DOCUMENT RESUME

ED 330 216 FL 019 117

AUTHOR Ramirez, J. David; And Others
TITLE Longitudinal Study of Structured English Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children. Final Report. Volumes 1 and 2.
INSTITUTION Aguirre International, San Mateo, CA.
SPONS AGENCY Department of Education, Washington, DC.
PUB DATE Feb 91
CONTRACT 300-87-0156
NOTE 1,444p.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Statistical Data (110)

EDRS PRICE MF12/PC58 Plus Postage.
DESCRIPTORS *Bilingual Education Programs; Comparative Analysis; Elementary Secondary Education; *English (Second Language); *Immersion Programs; Instructional Effectiveness; Limited English Speaking; Longitudinal Studies; Program Design; *Transitional Programs
IDENTIFIERS *Language Minorities

ABSTRACT

This study, undertaken from 1983-84 through 1988-89, compared the relative effectiveness of two alternative programs (structured English immersion and late-exit transitional bilingual education) with that of the program typically funded through the Bilingual Education Act, the early-exit transitional bilingual education program. The characteristics of each instructional program and the success with which each meets the needs of limited-English-proficient students are detailed. The report is in two volumes. The first volume describes the design and methodology of the study. An introductory section gives an overview of language learning options for limited-English-proficient students, and subsequent sections discuss data collection, data analysis, and characteristics of the programs examined and the participating students. The second volume presents the study's analyses and results, including conclusions arrived at concerning the relative effectiveness of immersion and the traditional program type in one- and two-program schools, intra-program analyses of late-exit transitional programs, comparison of programs, districts, and the norming population, and conclusions. Both volumes contain considerable data in tabular form and substantial lists of references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

FL

AGUIRRE
INTERNATIONAL

ED330216

FINAL REPORT: LONGITUDINAL STUDY OF STRUCTURED ENGLISH IMMERSION STRATEGY, EARLY-EXIT AND LATE-EXIT TRANSITIONAL BILINGUAL EDUCATION PROGRAMS FOR LANGUAGE-MINORITY CHILDREN

VOLUME I

Contract No.
300-87-0156

J. David Ramirez
Sandra D. Yuen
Dena R. Ramey
David J. Pasta

February 1991

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Submitted to:

U.S. Department of Education
Washington, D.C.

Aguirre International
411 Borel Avenue, Suite 402
San Mateo, CA 94402
(415) 349-1842

FL019117

ERIC

FINAL REPORT:
LONGITUDINAL STUDY OF IMMERSION STRATEGY, EARLY-EXIT AND LATE-EXIT TRANSITIONAL BILINGUAL EDUCATION PROGRAMS FOR LANGUAGE-MINORITY CHILDREN

VOLUME I

Contract No.
300-87-0156

J. David Ramirez
Sandra D. Yuen
Dena R. Ramey
David J. Pasta

February 1991

Submitted to:

U.S. Department of Education
Washington, D.C.

Aguirre International
411 Borel Ave., Suite 402
San Mateo, CA 94402
(415) 349-1842

The research reported herein was performed pursuant to Contract No. 300-87-0156 with the U.S. Department of Education. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official U.S. Department of Education position or policy.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

ACKNOWLEDGEMENTS

Aguirre International gratefully acknowledges the assistance of Drs. Keith Baker and Ricky Takai, Project Officers of the Office of Planning, Budget, and Evaluation, U.S. Department of Education, and the members of the study's Technical Advisory Panel who supplied constructive guidance at key stages during the study. We also wish to acknowledge the assistance of Dr. Barbara Merino, Professor, University of California at Davis, and Drs. Michael Long and Craig Chaudron, Professors, University of Hawaii, for the development of the classroom observation procedures. Dr. Steve Raudenbush, Professor, University of Michigan, Dr. John Willett, Professor, Harvard University, and Dr. Anthony Bryk, University of Chicago, were invaluable in assisting us with implementation of the Hierarchical Linear Model. Dr. David K. Billings, DMA Corporation, provided critical programming and technical support and managed myriad analytical tasks with grace and good humor.

We would also like to thank the district superintendents, program directors, principals, administrative personnel, and especially the teachers who were so cooperative in arranging schedules, completing questionnaires, permitting interviews, and access to records. The quality of any evaluation rests on the quality of the data. Local school personnel were outstanding in their efforts and support to help us realize the study's goals. This cooperation is greatly appreciated.

Members of the Technical Advisory Panel

Dr. Wallace Lambert
Professor, McGill University
Montreal, Canada

Dr. Eunice Lear
Project Director
San Diego Unified School District
San Diego, CA

Dr. Robert Linn
Professor, University of Illinois
Champaign, IL

Dr. Eva Pena-Hughes
Project Director
McAllen Independent School District
McAllen, TX

Dr. Merrill Swain
Director, Modern Language Center
Ontario Institute for Studies in Education
Toronto, Ontario, Canada

Dr. Gloria Zamora
Project Director, I.D.R.A.
San Antonio, TX

We would like to acknowledge the herculean efforts of project staff members without whom this challenging study could not have been effected:

Field staff

Rolando Castillo	Site Coordinator
Yolando De Duran	Site Coordinator
Carmen Dinos	Site Coordinator
Rosemary Fonseca	Site Coordinator
Teresa Hermosillo	Assistant Site Coordinator
Mary Jo Karam	Assistant Site Coordinator
Antonia Lacen	Assistant Site Coordinator
Carol Lay	Site Coordinator
Jorge Nakamoto	Site Coordinator
Vivian Rodriguez	Site Coordinator
Yolanda Suris	Site Coordinator
Estella Tristan	Assistant Site Coordinator
Martha Vallejo	Assistant Site Coordinator

Report preparation staff

Anne Abrams	Research Assistant
Teri Adams	Project Secretary
Hallie Dozier	Research Assistant
Zdena Valach	Research Assistant

DIRECTORY OF TABLES

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF FIGURES

I. LANGUAGE LEARNING OPTIONS FOR LIMITED-ENGLISH-PROFICIENT STUDENTS

Introduction

The primary task of the Longitudinal Study of Immersion Strategy, Early-exit and Late-exit Transitional Bilingual Education Programs for Language-Minority Children is to compare the relative effectiveness of two alternative programs (structured English immersion and late-exit transitional bilingual education) with that of the program typically funded through the Bilingual Education Act, the early-exit transitional bilingual education program. The characteristics of each instructional program is detailed as well as the success with which each program meets the needs of limited-English-proficient students. These data provide valuable information to policy makers and practitioners about alternative approaches for bilingual education and the requirements for the successful implementation of each.[1]

The Longitudinal Study of Immersion Strategy, Early-exit and Late-exit Transitional Bilingual Education Programs for Language-Minority Children is a six-year project begun in FY 1983-84 and ending in FY 1988-89. Year one of the project realized four major tasks: (a) finalizing the study design (Ramirez et al., 1984); (b) developing data collection instruments; (c) preparing literature reviews (Ramirez, Schinke-Llano, & Bloom, 1984; Schinke-Llano & Ramirez, 1984); and (d) selecting study sites (Ramirez, Wolfson, & Morales, 1985). Year two of the project resulted in the first of a four-year data collection effort producing information on the students in the study and their instructional programs (Ramirez et al., 1984). Years three and four of the study resulted in the Second and Third Year Reports wherein data

---

[1]Please see "Study Design of the Longitudinal Study of Immersion Programs for Language-Minority Children" for a description of the policy context for this study. The study, the history of federal bilingual education policy, and a summary of recent research and evaluation studies of instructional services to limited-English-proficient students are described in this report.

describing students and each instructional program collected were summarized and reported separately by year (Ramirez et al., 1987, 1988). These yearly reports allowed us to describe the instructional programs of the target students as they moved into the higher grades.

Purpose of the Report

The following "Final Report of the Longitudinal Study of Immersion Strategy, Early-exit and Late-exit Transitional Bilingual Education Programs for Language-Minority Children" describes the three instructional programs under study and the students participating in each. This report presents information on how structured English immersion strategy, early-exit, and late-exit transitional bilingual education programs were implemented at the study sites, and provides programmatic and contextual information to identify differences and similarities among these three programs. The report also describes the home and community backgrounds of the students participating in the study as well as their proficiency in oral language, reading, language arts, and math, as assessed in English and Spanish. In sum, implementation of each program is described vis a vis a common set of criteria, while achievement results are presented as a comparison of the relative effectiveness of the immersion strategy and late-exit programs in relation to that of the early-exit program.

The focus of this final report is to describe characteristics of the instructional treatments and identify: (a) similarities among class-rooms within a given instructional treatment, and (b) differences among the three instructional approaches. Ideally, similarity of classrooms within programs and differences between programs would suggest three distinct instructional treatments. Identifying such differences and similarities will help to determine how much differences in student achievement can be attributed to differences in instructional techniques.

Of equal importance, this report will also determine the extent to which students in the three instructional programs resemble each other, so that differences in student achievement in the separate programs can be attributed more clearly to differences in the programs and not to differences in critical student characteristics such as socioeconomic status.

Thus Volume I of the final report is concerned with answering the following questions: (a) Do we have three distinct instructional programs? (b) If yes, is each program consistent with its respective instructional model? And, Volume II of the final report addresses the following question: (c) What is the relative effectiveness of these three instructional programs?

To provide a context for this study, it is important to discuss briefly the history of federal bilingual education policy and summarize the recent research and evaluation studies of instructional services to LEP students. In the context of this discussion, the chapter will conclude with a listing of the study's objectives.

Historical Overview of Federal Policy in Bilingual Education to 1989

As more fully discussed in the "Study Design of the Longitudinal Study of Immersion Programs for Language-Minority Children" (Ramirez et al., 1984), federal involvement in education for language-minority children has focused almost exclusively on transitional bilingual education programs. In 1974, P.L. 93-380 provided the initial direction for this policy. It called for instruction in "the native language of the children of limited-English speaking ability...to the extent necessary to allow a child to progress effectively through the educational system." In 1978, the Bilingual Education Act (P.L. 95-961) expanded the definition of English proficiency from the earlier emphasis on speech to one focusing on reading, writing, speaking, and comprehension. While the revisions defined bilingual education as providing instruction in both English and the child's primary language, the goal of the educational program was narrowed. The earlier legislation (P.L. 93-380) had emphasized that the language-minority child progress "effectively through the education system." The new law specified that the instructional intent was to help the child "achieve competence in the English language."

Additional federal policies regarding bilingual education were established by (a) the Supreme Court's 1974 Lau vs. Nichols decision regarding discrimination in the San Francisco Unified School District; and (b) the resulting Lau remedies (Office of Civil Rights, 1975) which prescribed various types of bilingual education programs. While the original Lau decision did not prescribe any specific program or teaching strategies, the Lau remedies restricted federal funding under the subsequent 1978 Bilingual Education Act to programs that taught children in their native language until they were able to participate effectively in English-only, mainstream classrooms. As a result of these federal policies, federal education policy for language-minority children favored transitional bilingual education programs by funding through Title VII only those programs providing some primary-language instruction.

Curiously, although the Lau remedies never were adopted formally as regulations, their highly prescriptive requirements were used to determine which programs would receive federal funding. Central to the Lau remedies controversy was their conflict with many people's strong convictions about what constituted appropriate (and affordable) solutions to the problems of limited English proficiency. A particular concern was the assertion that it was not enough to provide an English-as-a-second-language (ESL) program because students were not able to study other subjects until they learned English. The argument behind the Lau remedies held that equal access (i.e., opportunity to study subjects other than English) could be ensured only if language-minority children were taught in their primary languages.

The 1978 Amendments to the Bilingual Education Act of 1968 provided further impetus for the use of the students' primary language in teaching language-minority children. Although the stated goal of the programs was to help language-minority children learn English as quickly as possible, the 1978 Amendments allowed use of the primary language "to the extent necessary" to ensure a meaningful education for language-minority children and a successful transition into mainstream programs. While the Amendments did not specify how the primary language was to be used, the only programs funded were those that included primary language instruction (i.e., met the Lau remedies' criteria). While instructional programs which did not use the primary language could not be funded under Title VII, English-only instructional services to limited-English-proficient students were provided through other funding sources, such as Chapter 1, Chapter 1 Migrant Education, state or local funds. However, little was known about their effectiveness.

The 1983 amendments to the Bilingual Education Act of 1968 (PL 98-951), signaled a major shift from the Lau remedies criteria. For the first time, the new amendments allowed funding of programs which did not require the use of the LEP child's primary language. These alternative models to primary language instruction primarily use English for instruction of LEP students. The child's primary language is used, if at all, only for support on an as-needed basis. The English is "tailored" or "sheltered" to accommodate the English proficiency level of the LEP students. English is adjusted in terms of pace (slower), simple rather than complex sentence structure, shorter rather than longer sentences, vocabulary, and use of realia (use of pictures, mime, etc., which would convey the meaning of the words being used). As this amendment heralded a major departure from prior policy, Congress limited the amount of Title VII funds that could be allocated to these alternative programs (and thereby the number of such alternative programs) to no more than ten percent of the total Title VII allocation.

The bilingual Education Act of 1968 was again amended on April 27, 1988 (PL 100-297). These amendments raised the ceiling on the funding of alternative programs from ten to twenty-five percent of t e total Title VII allocation. While the 1988 amendments also allo 1 the funding of developmental primary language instructional programs .e., where substantial use of the child's primary language), few, if y, were funded. The 1988 amendments, also for the first time, limited he number of years that a LEP student could be served by Title VII fu ds to no more than three, regardless of his/her proficiency in English.

In sum, in spite of the increased funding of special English-only programs since 1983, little if nothing is known about their effectiveness in meeting the educational needs of LEP students.

Characteristics of Programs

Four Methods for Teaching English to Limited-English-Speaking Students

Transitional bilingual education, however, is not the only means of teaching English to limited-English-speaking children in American public schools. Rather, four general alternatives have evolved to serve language-minority children: (a) submersion; (b) English as a second language (ESL); (c) transitional bilingual education (TBE); and (d) structured English immersion strategy. These programs differ in five main areas: (a) whether traditional all-English instruction is used; (b) whether special instruction in English as a subject is provided; (c) whether English is taught through the teaching of other subjects; (d) whether nonlanguage subjects such as mathematics are taught in the primary language of the limited-English-proficient students; and (e) whether the primary language of limited-English-proficient students is used (to supplement instruction in English). The following describes the four programs in terms of these characteristics (see Table 1).

Table 1

Characteristics of Programs Serving Language-Minority Children in the United States

Characteristics	Submersion	English as a Second Language	Transitional Bilingual Education	Structured English Immersion Strategy
Traditional all-English instruction used	Yes	Yes	No	No
Special instruction in English as a subject is provided	No	Yes	Yes	No
English is taught through the teaching of other subjects	No	No	Yes	Yes
Informal use of LEP student's primary language for clarification	No	No	Yes	Yes
Non-language subjects are taught in LEP student's primary language	No	No	Yes	No

The <u>submersion</u> approach is typical of the majority of services provided to language-minority students (Office of Bilingual Bicultural Education, 1981). In submersion programs, language-minority children are placed in ordinary mainstream classrooms where only English is spoken. No special provisions are made in these mainstream classrooms to help them learn quickly the English skills they need to succeed in school. Instead, they study the regular curriculum and are expected to perform as best they can. To the extent that the mainstream classroom represents an instructional program wherein English and other academic skills are developed, it can be considered as an instructional alternative to transitional bilingual education (i.e., one in which no special instructional adjustments are made). This alternative has been described as "sink or swim." Placement in submersion classes usually results when there are insufficient numbers of limited-English-proficient students in the same grade and school to "trigger" a bilingual classroom. Many of these students receive English as a second language instruction.

In typical English as a second language (ESL) programs, language-minority students spend most of their day in a submersion classroom, but receive some extra instruction in English. This special instruction is developed specifically to teach English as a second language. For subjects other than English, the language-minority students study the school's standard curriculum in English-only classrooms. In teaching ESL, the teacher may or may not use the primary language of the limited-English-proficient students. The California State Department of Education conducted a recent survey which determined that after submersion programs, ESL is the next most widely used instructional method for teaching language-minority students (Office of Bilingual Bicultural Education, 1981).

In transitional bilingual education (TBE) programs, language-minority students study subject matter in their primary language until they have learned enough English to succeed in English-only mainstream classrooms. Children in TBE programs generally learn to read first in their non-English home language and then in English. ESL often is used as a supplement to reduce the time needed to learn English. TBE programs are similar to submersion and ESL in that English usually is taught as a separate subject, but differ in that other content areas, including reading, are taught in the child's non-English home language, at least in principle.

The structured English immersion strategy program (SEISP) is proposed as an alternative to ESL and TBE programs. It is based on the results of Canadian French immersion programs for language-majority (i.e., English) speakers. While similar to ESL and some TBE programs in many ways, the SEISP also differs from them substantially. All instruction is in English. Rather than teaching English strictly as a subject, however, the SEISP endeavors to teach English through the various content areas. Prior knowledge of English is not assumed. Instead, teachers in SEISPs carefully tailor their English to a level the limited-English-proficient students can understand. For example, a lesson about a particular science concept also would teach the development and use of specific English language skills (such as vocabulary). Content thus becomes the medium for teaching language. Such teaching differs from transitional bilingual programs in that SEISPs present the subject matter exclusively in English, while TBE programs teach content in the students' primary language until they have learned English.

The SEISP teacher is bilingual and speaks the students' non-English home language. The student may use their primary language among themselves and to address the teacher. Generally, however, the teacher speaks to the students in English, using the home language only occasionally to provide or clarify instructions. Understanding the children's home language, the teacher can determine whether a child's difficulty with a given task stems from a problem with the language or with the content itself. As the child's primary language is not used formally for instruction, this type of instruction would be best described as a "sheltered English" program, rather than as a bilingual program.

Encouragement for Immersion: The Canadian Experience

Recent work in Canada has demonstrated that the immersion programs which use the second language exclusively can be very effective in teaching language-majority students a second language. It is the success of these Canadian programs that has stimulated interest in the usefulness of this approach for language-minority children in the United States.

There are major differences between Canadian immersion programs and U.S. structured English immersion strategy programs (see Table 2). In Canadian programs, language-majority students are grouped homogeneously by their proficiency in the second language, which is usually none; while a single class of minority-language LEP children in the American program may represent a broad range of second language proficiencies.

Canadian immersion programs and U.S. programs also differ in their criteria for success. The stated goal of Canadian immersion programs is full bilingualism, with near native-like competence the measure of success. The U.S. structured English strategy programs, on the other hand, concentrate on English proficiency as measured by the student's ability to function in a mainstream English-only classroom, and little or no concern is shown for the development of the student's first language. This minimal English competency level is far from native-like proficiency, in that it usually reflects social language rather than academic language skills.

Canadian teachers in immersion programs appear to adjust their language to L2 learners. However, the nature of these adjustments has not been documented as yet. In U.S. structured English immersion strategy programs, it is not known whether or to what extent teachers adjust their language for their L2 learners. The differences between the Canadian immersion programs and the U.S. structured English immersion strategy programs are found in Table 2. In sum, the Canadian and U.S. programs differ with respect to their goals (bilingualism vs. development of L2 only), target population (language majority vs. language minority), methodology (development of L1 vs. L1 not developed), and parent involvement or lack of involvement. As these differences between the Canadian and U.S. programs are substantial, there is a need to differentiate the names of the two programmatic efforts to avoid confusing them. As the Canadian programs were developed first, they will be referred to in this report simply as immersion programs. In contrast, the U.S. efforts are referred to as structured English immersion strategy programs in an effort to highlight the concern over and focus on those instructional strategies which might be used to teach English through the content areas.

ERIC

Table 2

Comparison of Canadian Immersion Program,
U.S. Immersion Strategy Program,
Early-Exit and Late-Exit Transitional Bilingual Education
Program Characteristics

Characteristic	Canadian Immersion Programs*	U.S. Structured English Immersion Strategy Programs	Early-Exit	Late-Exit
Bilingualism is a major program goal	Yes	No	No	No
(a) Develop L1	Yes	No	No	Yes
(b) Develop L2	Yes	Yes	Yes	Yes
Program is for language-minority children	No (language-majority)	Yes	Yes	Yes
Provides primary language instruction	Yes	No	Yes (limited)	Yes
Parents involved in developing program	Yes	No	Yes (required)	Yes
Teachers are bilingual	Yes	Yes	Yes	Yes
Students are from low SES homes	No (middle-class homes)	Yes	Yes	Yes
Students can use L1	Yes (only initially)	Yes	Yes	Yes
Subject matter is taught through L2	Yes	Yes	Yes	Yes
L2 develops as a function of learning subject matter	Yes	Yes	Yes	Yes

*These characteristics apply to the initial St. Lambert experiment.

At present, little is known about the existence or effectiveness of U.S. structured English immersion strategy programs with language-minority children. The "Description and Longitudinal Study of Immersion Strategy, Early-exit, and Late-exit Transitional Bilingual Education Programs for Language-Minority Children" was funded to provide this information.

Current Research and Evaluation Studies of Instructional Services to Limited-English-Proficient Students

The basic question addressed by this study is to determine the relative effectiveness of structured English immersion strategy programs and of two types of primary language transitional bilingual educational programs—late-exit and early-exit. The following highlights some of the more important and relevant studies to date of each approach.

Structured English Immersion Strategy Programs

Program organization. Canadian Immersion programs can be categorized along three dimensions: the "degree of immersion," the grade level of implementation, and the number of target languages included. With respect to the degree of immersion, programs are classified as either total or partial immersion. In early total immersion, the second, or target language is used for all of the curriculum from the beginning of the program (Genesee, 1984; Lapkin & Cummins, 1984). The students' native language is introduced into the curriculum after a period of time, generally after two, three, or even four years (Genesee, 1978; Genesee & Lambert, 1983; Lambert & Tucker, 1972). Once the native language is included in the program, its use as a medium of instruction may vary from 20% of the time (Morrison, 1981) to 60% (Genesee, 1978), depending upon the individual program or upon the particular grade level within a program.

Partial immersion, on the other hand, is characterized by the second language being used for less than 100% of the curriculum at the beginning of the program. Generally the second language is employed 50% of the time with decreases in use, if any, occurring only after a number of years (Genesee, 1984; Lapkin & Cummins, 1984).

With respect to the grade level of implementation, immersion programs can be described as early, delayed, or late (Genesee, 1984; Lapkin & Cummins, 1984). In early immersion programs, use of the second language begins in kindergarten or first grade. In delayed immersion programs, the second language usually is not introduced as a medium of instruction until the fourth or fifth grade. In late immersion programs, the second language is not used to teach content subjects until late in the elementary school years, or even early in the secondary school years. Both delayed and late immersion programs may be preceded by one or several years of traditional second language instruction, i.e., classes in which the second language is the subject--rather than the medium of instruction.

The grade level of implementation of an immersion program determines, among other things, one very important aspect of a child's education, namely the language in which literacy skills are initiated. In an early total immersion program, for example, literacy training is begun in the second language; in an early partial program, literacy is developed simultaneously in the native and second languages. Both delayed and late programs, on the other hand, allow for the establishment of literacy skills in the native language of the students after the second language skills are strongly developed.

Regarding the number of target languages included in the curriculum, programs are described as either single or double immersion. This designation obviously depends upon the inclusion of one or two second languages for instructional purposes. Hypothetically, of course, there could be triple immersion programs, quadruple programs, and so on. Theoretically speaking, all possible combinations of program types discussed can exist (e.g., early partial single immersion, delayed total double immersion, etc.), but single immersion programs predominate. While early programs are either total or partial, there is a tendency for delayed and late programs (e.g., usually everything is in French except for English language arts) to be partial in that instruction is provided in both languages (Genesee, 1984; Lapkin & Cummins, 1984).

Regardless of the degree of immersion, the grade level of implementation, or the number of target languages involved, all Canadian immersion programs have one essential characteristic: the second language is used not merely as the subject of instruction, but also, and more importantly, as the medium of instruction in subject matter classes. It is precisely this use of the second language in communicative contexts similar to those in which a first language is acquired that proponents say not only allows for, but also facilitates the acquisition of a second language.

Canadian Immersion Programs: Research Results

Since the inception of immersion programs in Canada, researchers have been diligently documenting their linguistic, cognitive, and social effects on the participating students. The Canadian findings have been used to advocate alternatives to TBE. It is important, however, to examine the studies in light of the four expressed program goals—first and second language competence, academic achievement, and psycho-social development.

First language development. Does participation in an immersion program retard students' development of their first language? No. Participation in an immersion program did not retard the development of Canadian students' ability to speak their primary language. Students in an immersion program spoke their primary language as well as their peers in non-immersion classrooms, at least during the early months (Swain, 1984). This undoubtedly is due to the pervasive presence of English in the school, community, and home environments. Regarding literacy-related skills, however, the picture is somewhat different. Early total immersion students initially are behind their non-immersion counter-parts. Yet within a year of the introduction of English language arts, immersion students perform as well on standardized

ERIC

English achievement tests as do the comparison students (Genesee, 1978). Second and third grade early partial immersion students sometimes perform less well on certain English literacy-related skills than do their English-program peers (Barik, Swain, & Nwanunobi, 1977). One possible explanation is that the simultaneous teaching of literacy skills that is done in early partial programs causes confusion for a period of time (Swain, 1984). If so, it is preferable to teach initial literacy skills in only one language. Further, according to Cummins (1981) and Swain (1983), if a minority-language group is involved, literacy skills should be taught in the children's first language, which is probably the more highly developed.

Other studies of the effect of an immersion program on first language development reveal valuable information as well. Genesee (1974), in a study of the writing in English (first language) of fourth-grade immersion students, found that the immersion group scored lower than the comparison group on spelling, but higher on measures of creativity. In a study of the global assessment of fifth graders' compositions in their first language, Lapkin (1982) found no difference between experimental and control groups. In a parental survey conducted by McEachern (1980), 80% of the parents with children in immersion programs felt that their children were experiencing no problems in English communication (i.e., first language). Finally, children in kindergarten, first grade, and second grade immersion programs have been judged superior to non-immersion students on measures of communicative effectiveness (Genesee et al., 1975).

In sum, there are no long term English language deficits resulting from participation in Canadian immersion programs, although sometimes there are short term lags in the development of literacy related skills.

Academic achievement. Studies show that participation in an Canadian immersion program develops strong academic and cognitive skills. Swain and Lapkin (1982) reviewed standardized mathematics tests of early total immersion students, grades one through eight. On average, the students scored as well as or better than their non-immersion counterparts on 35 of the 38 tests administered. Similarly, on 14 administrations of a standardized science test to students in grades five through eight, early total immersion students and non-immersion students scored equally well. These positive findings are not entirely replicated in immersion programs that were not early total immersion. Barik and Swain (1977) report inferior mathematics scores for early partial immersion students beginning in the third grade. In addition, Barik and Swain (1978) found inferior performance by early partial immersion students in science beginning at the fifth grade level. Among students in late immersion programs, Barik and Swain (1976) observed occasional inferior performance in science when the experimental groups had received only one or two years' instruction in French as a second language (FSL) prior to beginning the immersion program; however, these lags disappeared in subsequent years. A similar phenomenon was observed in mathematics performance (Barik, Swain, & Gaudino, 1976).[2] On the other hand, in instances where late immersion students had received FSL instruction since kindergarten prior to their program entry, their performance in content area subjects was equivalent to that of the comparison group (Genesee, Polich, & Stanley, 1977). This has been found even when locally or systemwide achievement tests based on local curricula were used (Genesee, 1983).

While these are not related to specific content areas, several studies suggest that there may be cognitive benefits associated with bilingualism developed in immersion programs. In a seven-year study of immersion and non-immersion students matched for IQ and socioeconomic status, Scott (1973, cited in Lambert, 1984) found that the fifth and

---

[2]In all such studies of academic achievement, Swain (1984) cautious that the language of test administration can affect students' performance. In other words, even though the second language is used for instruction, the first language may be the more appropriate one for assessing content knowledge.

sixth grade immersion students scored higher on divergent thinking, one measure of cognitive flexibility. Further, Barik and Swain (1976) and Cummins (1975, 1976) found increases in immersion students' IQs or in divergent thinking that were not found in the comparison groups. Finally, students whose IQs are below average or who have learning disabilities are not at any more of a disadvantage in immersion programs than they are in all-English programs (Bruck, 1979; Genesee, 1976; Swain, 1975). In fact, Bruck (1978) suggests that, at least with respect to French language acquisition, learning disabled students in immersion programs may have an advantage in that they acquire a second language. While these studies are indeed significant, there is no doubt that much research remains to be done on the relationship between bilingualism and cognitive processes.

Psychological and social effects. Numerous benefits from immersion programs have been documented in the areas of psychological and social development. Lambert and Tucker (1972), for example, found that immersion students have more positive attitudes towards French Canadians than do their non-immersion English-Canadian peers, although such differences are less evident in higher grades (Genesee, 1984). Cziko, Lambert, and Gutter (1979) report that immersion programs appear to reduce English Canadians' perception of the social distance between themselves and French Canadians. Fifth and sixth grade immersion students, when asked to write a composition on why they liked being Canadian, more frequently mentioned the linguistic and cultural diversity of Canada. Non-immersion students tended to cite the natural beauty of the country (Swain, 1980). Clearly, then, the goal of increased cultural understanding appears to be a by-product of immersion programs. Alternatively, Genesee (1983) has proposed that immersion students' second language proficiency reaches a plateau relatively early in the program and does not continue to develop because the communicative demands made upon them in class do not expand (Genesee, 1987).

ERIC

Second language development. Ironically, one of the most important goals of immersion programs, that of achieving competence in the second language, appears to be the most problematic. For example, when the French performance of early total immersion students is compared with that of non-immersion students who study French as a second language, immersion students are consistently superior (Swain, 1984). However, when compared with native speakers of French, the immersion groups appear to need six or seven years to achieve average performance in the receptive skills of listening and reading (Swain & Lapkin, 1982). Further, with respect to production skills (speaking and writing), immersion students have not shown native-like proficiency (Genesee, 1978; Harley, 1979, 1982; Spilka, 1976). Plann (1976), in a United States' study, posits the development of a classroom dialect peculiar to immersion programs which is fostered by the students' reinforcement of incorrect usage and which in fact may work against students' developing native-like proficiency.

All of the studies cited with the exception of the last, are Canadian. Few studies on immersion programs in the United States exist, first because of a paucity of programs, and second because of a lack of financial support for such research. The Culver City Spanish Immersion Program, however, because of its association with UCLA, has been evaluated. Findings replicate those of Canadian programs (Campbell, 1984). Academically, immersion students have performed either as well as or better than their non-immersion peers. Their English skills are equivalent, with the exception of mechanics and spelling. Some attitudinal improvement is evident. Their Spanish, however, while competent, is not native-like. Thus, the original goal of native-like competency in Spanish oral language skills has been re-examined.

All of the research studies cited here concern themselves with the performance of language-majority students, those for whom participation in the program does not threaten the maintenance of either their home language or home culture. The obvious unanswered research question,

ERIC

then, is whether such language immersion programs can be adapted for use with language-minority students. While no empirical evidence exists to support or refute the notion, researchers in both Canada and the United States have been quick to caution against such an attempt (Genesee, 1984; Hernandez-Chavez, 1984; Lambert, 1984; Swain, 1984; Tucker, 1980). Tucker (1980, p. 2), for example, emphatically states:

> We have not previously, and we will not in the future, recommend on the basis of these careful, critical, and longitudinal studies, that Mexican American, Franco American, or other non- or limited English speaking youngsters in the United States be submerged in English medium programs.

Genesee (1984, p. 53) echoes this position by warning:

> To recreate these conditions and implement these approaches to facilitate academic and language learning among minority language children will require more than changing the names of the languages involved. Indeed, this task may require changing the basic structure of immersion education as it is known to apply to majority language children.

This study will provide the first documentation of how effective the U.S. adaptation of the Canadian language immersion programs is in meeting the learning needs of limited-English-proficient students and how these programs compare with two types of primary language development programs.

Characteristics of Primary Language Programs

In general, primary language instructional programs in U.S. public schools occur within the context of transitional bilingual education programs. How the primary language is used, the amount of use, the context in which it is used, those using it, and how each of these factors varies and is combined with the others produce markedly different instructional programs. Moreover, primary language instruction is only one of numerous factors which interact to shape the bilingual program. To fully appreciate and understand primary language instructional programs, it is important to understand the context in which each occurs.

In defining bilingual education, Fishman asserts that "bilingual education implies some use of two (or more) languages of instruction in connection with teaching other than language per se" (cited in Trueba & Barnett-Mizrahi, 1979, p. 2). Clearly, in this view, neither foreign language education nor the limited use of a native language would constitute bilingual schooling.

A review of selected bilingual education models (Ramirez, Schinke-Llano & Bloom, 1984) reveals marked differences among researchers as to what they consider critical characteristics of bilingual programs. To account for all possible types of bilingual education, Mackey develops an elaborate classification system. At the heart of his taxonomy are language interactions; that is, the give and take in conversation between speakers. Paulston (1980) states that the following program characteristics must be observed to define the instructional program: the languages used in the classroom; how the languages are sequenced (i.e., when the second language is introduced); how much each is used, and for what purpose; the linguistic competencies and ethnicity of the classroom teachers; availability of instructional materials in each language; and the language of the student's home/community. Others are concerned with program goals and how well they match the social/political conditions of the community; that is, the extent to which there is a need for an individual to be bilingual in a given community (Fishman & Lovas, 1970; Kjolseth, 1972).

Research Findings: Primary Language Programs

Given the diversity of primary language programs in the United States, what, if anything, is known about their effectiveness? What information is available outside of the United States? Data are available from two sources, the United States and Canada.

U.S. research. Unlike the language immersion programs in Canada, which have been carefully monitored by researchers since their inception, bilingual education programs in the United States have not been the subject of such systematic, careful research or evaluation. The result of this lack of research is that little more is known about the efficacy of primary language programs for limited-English-proficient students than was known before their implementation 16 years ago.

Of the research studies and evaluation reports which are available, a large number suffer from problems of design, methodology, or analysis that render their findings virtually meaningless. Zappert and Cruz (1977) and Dulay and Burt (1978) report rejecting nearly 95% of the 38 studies and 175 reports they examined because of one or several of the following shortcomings:

- lack of control for socioeconomic status of subjects;
- lack of control for initial language proficiency of subjects;
- lack of control group or baseline data for comparison;
- sampling problems (e.g., inadequate sample size, non-random sampling, excessive attrition between pre- and post-testing);
- significant differences in teacher characteristics; and
- insufficient data or inappropriate statistical analyses.

The review by Zappert and Cruz (1977) has been criticized for its conclusion that the neutral results would suggest that the evidence does not show that the programs hurt the children, and therefore the programs should be supported. Given the cost and myriad administrative problems in mounting primary language programs, some argue that the benefits do not outweigh the problems. Moreover, the perceived benefit of speaking a language other than English goes beyond the goals of current public policy which are concerned with English proficiency and not primary language proficiency.

Of the remaining 12 studies and reports deemed methodologically sound by Zappert and Cruz (1977) and Dulay and Burt (1978), 58% showed positive results of bilingual treatment, 41% were neutral, and 1% were negative.

Three highly publicized reviews of the literature found evidence in support of the effectiveness of primary language instruction to be lacking. The first is Language, Ethnicity, and the Schools (Epstein, 1977) in which the author, after interviews with politicians and administrators, reports that no evidence of the educational effectiveness of bilingual programs exists. The second is he 1978 series of American Institutes for Research (AIR) reports which concludes that students in Title VII Spanish bilingual programs performed lower in English than their non-Title VII counterparts and equally in mathematics. The AIR study, however, has received severe criticism for having included large numbers of English-dominant and English monolingual students in experimental groups (Dulay & Burt, 1979). Finally, there is the Baker and de Kanter report (1981) commissioned by the U.S. Department of Education, which investigated the performance of bilingual program students in both English and mathematics. Baker and de Kanter found the evidence in support of transitional bilingual education programs to be inconclusive and exclusive reliance on this instructional method for federal educational policy to be unjustified. They argued that other alternatives should be investigated such as immersion strategy language programs for which they suggest there is some evidence of success. Like the AIR study, the Baker and de Kanter report has drawn strong criticism, in this case for its procedures in tallying study results, as well as in interpreting the results (Willig, 1981). Further, there is major concern over whether the immersion strategy program cited as successful is truly an immersion strategy program.

Nonetheless, one cannot overlook Baker and de Kanter's statement that, "We know that TBE works in some places and fails in others, but we do not know why and therefore cannot specify in what situations TBE should or should not be used" (1983, p. 53).

Interestingly, each reviewer cited confirmed the lack of sound research on the effectiveness of primary language programs. A recent major study, however, has investigated bilingual education from a different perspective from that utilized in the studies just discussed. The "Significant Bilingual Instructional Features" (SBIF) study, funded by the National Institute of Education, had as its goal the identification, description, and verification of important features of bilingual education for the instruction of LEP students (Tikunoff, 1983a). Both quantitative and qualitative procedures were used to analyze:

- o organization of instruction;
- o allocation of time;
- o language use;
- o teaching behaviors;
- o academic learning time (ALT);
- o student participation; and
- o classroom, school, community, and context variables.

Successful bilingual teachers (so designated through nominations by administrators, other teachers, parents, and students) were observed. Five significant features of instruction were identified:

- o congruence between instructional intent, organization and delivery of instruction, and student consequences;
- o use of active teaching behaviors (e.g., communication, engagement, monitoring, and feedback);
- o use of L1 and L2 for instruction;
- o the use of English as a medium of instruction; and
- o use of information from the LEP students' home culture.

The first two instructional features, of course, are characteristic of effective teaching in general. The last three, however, relate directly to the instruction of LEP students. In addition, it was found that instructors used English 70% of the time during basic skills instruction, while they used the students' home language approximately 30% of the time. L1 was used most often by teachers to clarify lesson content being given in L2. Further, substantial use of L1 was associated with positive learning behaviors on the part of LEP students. The use of the target language for instruction is consistent with Canadian immersion research. Not surprisingly, given the variety of programs identified in previous studies, the SBIF study found "the implemented form of bilingual education to be complex, diverse, and frequently influenced by conditions external to the classrooms" (Fisher & Guthrie, 1983, p. iii).

Despite the new perspective which the SBIF study brings to the analysis of bilingual instruction, certain criticisms have been leveled, and crucial questions remain unanswered. The most criticized aspect of the SBIF study is the use of academic learning time (ALT) as an objective measure of student participation (Cazden, 1984; Cummins, 1983; Paulston, 1983). ALT, as defined by Tikunoff (1983b), is the amount of time a student spends in a particular content area engaged in learning tasks with a high degree of accuracy. Cazden (1984), however, points out that tasks which yield accurate responses are not necessarily the most intellectually relevant. Cummins (1984) warns that ALT is a necessary, but not sufficient, index of effective teaching. Further, Paulston (1983), highlighting the acknowledged role of error in second language acquisition, states, "Student accuracy rate as an aspect of ALT is not valid for the process of language acquisition" (p. 70).

The language use pattern observed in the SBIF study also is problematic. The validity of the language use patterns is questionable as no comparisons were made with classrooms without these patterns. Despite the fact that the successful teachers used English approximately 70% of the time for basic skills instruction and changed to the students' first language for clarification of concepts and instructions, researchers warn that it is neither logical nor prudent to conclude that this pattern is necessarily beneficial for second language acquisition (Paulston, 1983; Wong-Fillmore, 1983). While SBIF researchers (Tikunoff, 1983b) are careful to distinguish between code-switching (which continues the communication) and code alternation (which spontaneously repeats or clarifies the message), Paulston (1983), in particular, warns against prematurely recommending such code alternation as an effective technique for bilingual teachers.

Wong-Fillmore (1983), citing the role of modified input in second language acquisition, observes that teachers in the study change to L1 rather than modify their input in English. She asks specifically, "How well did the English used in these classes work as input? What were the effects of the L1 alternations on the English that was used in lessons?" (p. 53). Finally, McLaughlin (1982) in his literature review, states:

> It would be premature to regard the issue (use of the first language) as settled. Most likely, decisions as to when and to what extent each of the bilingual child's two languages should be used in the classroom depends on social, psychological, and linguistic factors. Some children, in some circumstances, need more support in their first language than others do. (p. 34)

Canadian research results supporting bilingual education. Interestingly, the research documenting the effectiveness of the Canadian foreign immersion program provides the strongest suggestion for the effectiveness of a primary language instructional program. It is important to remember that the child's primary language and the target language are both developed and used for instruction in the Canadian immersion programs. Different immersion program models have been developed in Canada. They differ in two factors: (a) when the second language is introduced (i.e., early = kindergarten to grade three, delayed = begin in grade four or five, or late = begin in grade seven or eight), and (b) the proportion of instructional time for which each language is used (total or partial). A critical feature is that the languages are separated by subject or by time. For example, social studies always may be taught in the second language and math in the first language. Or, the second language is used only in the morning, and the first language in the afternoon. Or, the languages are used alternately every other day or week. Both are used, but their use clearly is differentiated.

The use of both the target and the primary language for instruction as well as the goal of bilingualism make Canadian immersion programs comparable to U.S. late-exit bilingual programs.

The effectiveness of the Canadian immersion programs has been amply documented (Genesee, 1987; Lambert & Tucker, 1972). Students in all types of Canadian immersion programs consistently have outperformed in second language skills (speaking, reading, and writing) their counterparts in regular monolingual programs. Genesee notes that late Canadian immersion programs have been found to be very successful:

> the effectiveness of the late option has been attributed to the cognitive and linguistic maturity of older students which are thought to contribute to more rapid progress in second language learners. This option may be of potential interest to American educators and parents who wish to concentrate on first language development prior to extensive exposure to a second language. (Genesee, 1985; p. 35)

This would be consistent with Cummins (1981) who suggests that proficiency in the primary language will facilitate the transfer of skills from the child's primary language to English, thereby improving bilingual proficiency.

Beyond these data, there is little evidence substantiating the effectiveness of primary language programs. In summary, as with

ERIC

structured English immersion strategy programs, much remains to be learned about the critical elements of a bilingual primary language programs and how effective they are in meeting the needs of limited-English-proficient students. This study describes the characteristics and effectiveness of two transitional bilingual education programs: early-exit and late-exit. This information will contribute to a better understanding of the nature and efficacy of primary language programs.

While the program definitions above provide a general outline of how each program functions, they do not operationalize how languages should be used in the classroom. Many practitioners and researchers of bilingual education in the United States assert that, in practice, these programs as implemented are best defined administratively in fiscal terms, as programs receiving a certain type of funding. They feel that, in fact, many programs seldom if ever use the primary language of the children they serve (Wong-Fillmore et al., 1983). Even within programs that are to use two languages, there is substantial program variation in the patterns of use of each language (Legarreta, 1977). However, this perception stems more from anecdotal observations or limited data rather than from large scale studies of bilingual classrooms.

Classroom Language Use For Instruction And Effective Second Language Teaching Behaviors

Research on language use patterns and effective teaching behaviors in second language programs has shifted from simplistic comparisons of students' achievement under different treatments to more complex studies incorporating documentation of fidelity of treatment (Ramirez & Merino, 1989). This change reflects recognition of a basic tenet of educational research that program effects can be understood only after the program treatment has been defined operationally and observed systematically to confirm fidelity of treatment implementation (Baker & de Kanter, 1981; Willig, 1985; Wong-Fillmore & Valadez, 1986). To date, few studies on bilingual education have included classroom observational data to confirm program implementation. As noted by Ramirez and Merino (1989), those few studies that have included such data, have drawn from teacher effectiveness literature (Dunkin & Biddle, 1974; Long, 1983). From this literature, four observational approaches have been used: (1) process studies which attempt to describe the process in which the two languages are used with bilingual children (Schultz, 1975; Milk, 1980); (2) process/context studies which describe the relationship of process to context, for example, the distribution of language use in different program models (Legarreta, 1977); (3) process/process studies which describe the relationship of process to process, for example, how certain teacher behaviors such as feedback can affect the responses of students (Chaudron, 1977); and (4) process/product studies which describe the relationship of process to product, in which effective teaching behaviors are identified in relationship to language use and their effect in promoting student achievement (Ramirez & Stromquist, 1979; Legarreta, 1977).

Process studies represent much of the early observational work on bilingual classrooms. Case studies were effected to describe how language was used in one classroom or program. Using ethnographic techniques to describe the nature of classroom discourse in a Boston bilingual classroom (Trueba & Wright, 1981), Schultz (1975) found teachers tended to favor the use of English for instruction, using Spanish principally to control student behavior. Both students and teachers felt that Spanish should not be used. The most complex instruction occurred in English. Mackey (1972) used a similar approach to describe language use patterns in a Berlin school. While the results of case studies are not generalizable, their value is in the identification of critical questions and/or processes that should be examined through more intensive and controlled studies. The value of this contribution, however, is offset by the limited number of classrooms with which it can be used. The case study approach is very labor intensive, requiring a great deal of resources. This investment precludes the use of case studies with a large number of classrooms. Finally, the case study approach does not allow for establishing a

ERIC

clear relationship between observed classroom processes and student outcomes.

Legarreta (1977) was the first to systematically study the relationship of language use and program model (i.e., process/context study). Using real time coding and a category observation instrument, she found that when bilingual kindergarten teachers used a concurrent translation approach, they used English more than three-quarters (80%) of the time. It was only when teachers alternated the use of language by day (one day English only and one day Spanish only) that use of each language tended to be more or less equal. Other studies exploring the relationship between type of language used and program model examined specific language structures. Hayman and Tucker (1980) used a category observational scheme to code the language used by three French immersion teachers from two English schools and three teachers from two French schools in Montreal. They found no differences in the frequency of nine grammatical structures, nor did they find differences in the general teaching strategies used. They found that regardless of the program model, teachers have similar questioning, reinforcement and error correction patterns. Hayman and Tucker considered these findings and concluded that teacher classroom discourse is similar across program models, at least when it is analyzed broadly. These studies are important in that reliable observation procedures were used and differences were analyzed across programs. However, the small sample size used in each study severely limits their generalizability and makes it very difficult to differentiate teacher and program differences (Ramirez & Merino, 1989).

Process/product studies frequently related teachers' and students' language behaviors to an outcome measure, usually student achievement (Politzer, 1977). In one study 18 bilingual elementary teachers teaching English as a second language were recorded on videotape (Ramirez & Stromquist, 1979). Student gains were related to specific teacher behaviors. Student oral production was found to improve when teachers: (1) required students to manipulate concrete objects following a teacher command; (2) questioned students regarding information previously presented by the teacher; (3) explained the meaning of new words; (4) corrected students' grammatical errors directly by providing the correct structure; and (5) varied the type of teacher behaviors. In contrast, student oral production decreased when teachers modeled or corrected pronunciation errors. The value of process/product studies is that they directly relate specific instructional practices to student gains. However, as they focus on classroom behavior they tend to ignore individual differences in the learner, such as how motivation and cognitive style interact with the learning process. As these studies are correlational, results are limited to statements of co-occurrence rather than causality of student achievement (Ramirez & Merino, 1989).

Rating scales also have been used in process/product studies to record the frequency of classroom behaviors by categorizing them from low to high. Wong-Fillmore et al. (1985) used such a measure to analyze 17 bilingual and English-only classrooms. Third and fifth grade classrooms were observed during a period of one year through the use of video, audio, and live recordings. Student gains on oral language and achievement tests were used to identify successful and unsuccessful classrooms. Gains in English production skills were found to be related to: interactional opportunities (i.e., fair allocation of turns), quality of instructional language (e.g., contextualization of information), and quality of teaching (e.g., clarity of instructional goals). These variables were found to affect Chinese and Hispanic students differently. For example, while opportunities to interact with other students helped Hispanics, they were not helpful to Chinese students. The advantage of a rating approach is that it allows one to observe high inference and more complex behaviors. This procedure is limited in that it is subjective and inter-rater reliability may be difficult to achieve.

Both the Ramirez and Stromquist (1979) and Wong-Fillmore et al. (1985) studies suffered from a lack of a clear operational definition

ERIC

of an instructional model and classrooms strongly adhering to a specific model. Wong-Fillmore et al., for example, observed that teachers in the study used their students' primary language from zero to 24% of the time. Given this diversity in treatment, it is extremely difficult to identify effective bilingual classroom techniques (Ramirez & Merino, 1989).

Process/process studies represent a different line of research. This body of work attempted to understand teacher and student language use patterns in the classroom and how they affect one another. Gaies (1977) and Chaudron (1977) found that teachers adjust the complexity of their speech to ESL students to accommodate their students' English proficiency level. Increasing the wait time when asking second language learners a question was found to increase the number of correct responses (Holley & King, 1971, cited in Cohen, 1975). Schinke-Llano (1983) found that some bilingual elementary teachers interacted differently with fluent and limited-English-proficient students. For example, teachers provided fewer academically related utterances to limited-English-proficient students than they did to English-proficient students. Other research efforts examined the relationship of instructional strategies in ESL classrooms to student engagement (Nerenz & Knop, 1983; Tikunoff & Vazquez-Faria, 1982) or to student perception of effectiveness (Omaggio, 1982). The value of process/process studies is their goal of trying to understand the teaching process itself (i.e., nature of the treatment) as it occurs rather than its relationship to outcome. While this emphasis on process is critical to understanding the nature of the instructional treatment (i.e., what actually happened), it does not help us to understand the effect of treatment on achievement.

To ensure a comprehensive study, this study draws from work done in each of the four research approaches outlined above (process, process/context, process/outcome, and process/process studies). A clear definition of instructional treatment was provided for each of the three instructional programs. Fidelity of treatment implementation

ERIC

was effected through detailed documentation of teacher and student language behaviors as well as a plethora of background information on students, their parents, teachers, classroom, school site, and school district. These data were augmented by annual assessment of each student's achievement. In contrast to the limited sample of the studies cited, and those not cited, this is the first study of programs serving limited-English-proficient students to have gathered detailed information from over 2,000 children from over 800 classrooms over a period of four years. The scope of data will allow for a good understanding of treatment and of its relationship to student outcome.

Study Objective

The primary objective of this study is to assess the relative effectiveness of structured English immersion strategy, early-exit, and late-exit transitional bilingual education programs. This will be done by determining the extent to which the educational needs of limited-English-proficient students are addressed through a proposed instructional program that provides all formal instruction in English and/or one in which formal instruction is provided in the child's primary language. This comparison is important in that despite the positive outcomes of Canadian immersion programs for language-majority children, there currently is insufficient empirical evidence to argue for or against the use of an adaptation of the Canadian approach with minority language children. This will be accomplished by comparing the performance of language-minority children in structured English immersion strategy programs with their peers participating in late-exit and early-exit transitional bilingual education programs. To meet this objective, the study addressed the following evaluation questions:

- What is the relative effectiveness of structured English immersion strategy, early-exit, and late-exit transitional bilingual education programs in meeting the learning needs of limited-English-proficient students with respect to English language proficiency and non-language academic skills?

- o If there are differences, under what conditions (such as for which students, taught by what type of instructor, using what types of methodology)?

- o What are the characteristics of each program?

- o What generalizations can be made about the education of limited-English-proficient students?

Drawing from data collected across the four years of the study this report provides information on the homogeneity of the instructional program across classrooms within each instructional treatment, the differences and similarities in classrooms among treatments, and the comparability of student/community characteristics within and among instructional treatments. Finally, and very importantly, outcome data is reported separately for each nominal program as well as operational programs. Student achievement data are presented as trends.

Organization of the Report

The remainder of Volume I is organized roughly into three sections—methodology, results, and summary of findings. The first section, methodology, describes the data collection (Chapter II) and analytic procedures (Chapter III). In the second section, results of the characteristics of each program are described such as classroom language, course content, instructional staff and administrative context in which each program functions (Chapter IV); and student characteristics (Chapter V). The third section, Chapter VI, comprises the conclusion.

ERIC

II. DATA COLLECTION OVERVIEW

Introduction

The intent of this chapter is to provide an overview of the implementation of the study. Information is provided on how the specific instructional programs and students for participation in the study were selected; how the data collection was conducted; and the procedures used for data management and quality control.

Sample Selection

The scope of work for this study required that the relative effectiveness of three instructional programs be assessed: structured English immersion strategy, early-exit, and late-exit transitional bilingual programs. Only those programs serving Spanish-speaking limited-English-proficient students were considered. This limitation reflected the need to ensure a study design that allowed clear program comparisons with available resources. Inclusion of programs serving more than one language-minority group would have required a more complex design and much greater resources. The following section outlines the steps taken to identify study sites.

Study Sites

Immersion strategy programs. Four steps were followed to identify English immersion strategy and late-exit transitional bilingual programs for participation in the National Longitudinal Study of Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Programs for Language-Minority Children.[3] In Step 1, Aguirre International/SRA Technologies conducted a national telephone survey over a

---

[3]See Ramirez et al. (1985), "Description of Immersion Strategy Programs in the United States," for a complete report.

six-week period to identify potential programs. Personnel in state educational agencies were contacted, as were bilingual educators, members of bilingual program support agencies, educational administrators, and organizations. They were asked if they knew of elementary instructional programs serving limited-English-proficient (LEP) students which met the study criteria (see Table 3). These initial interviews provided referrals to both potential study sites and other individuals or agencies that might act as resources for information. As a result of the telephone survey, a total of 91 programs in 29 states as potential immersion strategy sites were identified. Of these, 31 were subsequently excluded because additional information from the initial contact person indicated that they did not meet selection criteria.

Table 3

Immersion Strategy Program
Telephone Survey Selection Criteria

1. Subject matter instruction is designed to teach English.
2. All formal instruction is in English.
3. The primary language is not used for formal instruction but is used on an informal, as needed basis.
4. Teacher is trained in language development (e.g., holding a bilingual or ESL credential).
5. Teacher is bilingual in the LEP students' primary language.
6. There are at least 60 children per grade level.
7. Project classrooms have LEP students who speak the same primary language.
8. There is a bilingual program available for comparison.

Step 2 consisted of contacting project directors or key program staff from each of the 60 remaining referred programs by telephone to verify the extent to which the selection criteria were met. This second survey revealed that, of the programs initially identified, 30 were essentially ESL programs for multi-language populations, 20 were early-exit transitional bilingual education programs, and one was an all-day ESL program. Only nine sites appeared to meet the study's definition of an immersion strategy program.

Step 3 required a site visit by study staff to the most promising sites to determine the extent to which identified sites met the study's selection criteria and to confirm program characteristics. Teachers, school administrators, and project directors were also interviewed. Although the classroom observations were brief, they were sufficient to allow the observation of: (a) the teachers and aides working with individual, small group, and whole class instruction; (b) the instructional materials; (c) the lesson plans; (d) the student assignments; (e) scheduling; and (f) classroom management. During the interviews, the question of how the immersion strategy program was implemented at each school site within classrooms, within grade levels, and if appropriate, across grade levels was asked. Descriptions of how the immersion strategy classrooms differed from the regular mainstream classrooms, how teachers used the lesson plans with LEP students, how they used Spanish in the classroom, how they grouped students for instruction, how they assigned homework, and what instructional materials and specialized in-service training were available were then compiled.

Finally, Step 4 consisted of a letter of invitation to those programs that most closely resembled the program model as defined by the study. An initial site selection was made based on the degree to which sites met the operational definition of immersion strategy programs and the criterion for geographic spread. The primary selection criteria for an immersion strategy program were: (a) the teacher uses English exclusively for instruction; (b) the teacher's use of Spanish is informal, such as explaining directions; (c) the teacher uses the content areas to teach English; (d) students are free to use

Spanish among themselves and with the teacher; and (e) the teacher is bilingual in English and Spanish.

Of the nine sites that were selected, six were from Texas, two from California, and one from Florida. Of these, four agreed to participate in the first year of the study (two from California and two from Texas). The remaining sites declined to participate; three cited extensive involvement in an ongoing evaluation, and two had already committed extensive teacher time to completely revising the district curriculum pursuant to recent state mandates.

Early-exit transitional bilingual education programs. Each district that had an English immersion strategy program also had an early-exit transitional bilingual education program. In general, all early-exit programs met the program characteristics listed in Table 4. To maximize comparability of students between immersion strategy and early-exit programs, included in the study was a sample of the early-exit classrooms from each of the four districts with immersion strategy programs.

Late-exit transitional bilingual education programs. The next step in the sample selection was to identify late-exit transitional bilingual education programs. To do this procedures similar to those used to identify immersion strategy programs were used. A national telephone survey and a follow-up survey were conducted and sites identified during this process were visited. The search confirmed that there were only five school districts in the country that met the criteria for a developmental primary language (i.e., late-exit) transitional bilingual education program (see Table 4) — two in California, one in Florida, one in New York, and one in Wisconsin. Three of these districts — California, Florida, and New York— participated in the study. Of the remaining two, one program was terminated by the school district at the end of Year 1 for lack of funding. In the other, district administrators agreed to participate, but instructional staff voted against participation in a study they perceived to be potentially misrepresented by the U.S. Department of Education. All districts having a late-exit program offered only that program alternative (except for the one district terminating its late-exit program). That is, they did not offer an early-exit or an English-only immersion strategy program. The late-exit transitional bilingual program represented a district-wide policy.

Table 4

Instructional Programs to be Examined by the Longitudinal Study of Immersion Strategy, Early- and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children

Immersion Strategy Program Characteristics

(a) Teacher uses L2 exclusively for instruction; teacher's use of L1 is informal, such as giving or clarifying directions.
(b) Content areas are used to teach L2.
(c) L2 is used to teach content.
(d) Students are free to use L1 among themselves and with teacher.
(e) Teacher is bilingual, speaking.
(f) Children are mainstreamed into English-only programs as soon as they have demonstrated proficiency in English. This transition into an English-only program usually occurs within two or three years after entry into the immersion strategy program.
(g) There is a limited primary language component.

Early-Exit Transitional Bilingual Education Program Characteristics

(a) Teacher is bilingual.
(b) Teacher uses both L1 and L2 for instruction.
(c) L1 language arts skills may be developed first, before introduction of L2 language arts, or at the same time.
(d) Instruction in L1 is minimal, not more than one hour a day.
(e) Use of L1 and L2 is not differentiated by teaching staff.
(f) Teachers using L2 have "native" or "near-native" L2 skills.
(g) Children are mainstreamed into English-only programs as soon as they have demonstrated proficiency in English. This transition into an English-only program usually occurs within two or three years after entry into the early-exit program.
(h) There is a limited primary language component.
(i) There is an L2 language arts component.
(j) Cultural sensitivity is reflected in the program by the teacher, the instructional materials, and the children's tasks.
(k) All content areas are taught in L2.

Table 4 (continued)

Late-Exit Transitional Bilingual Education Program Characteristics	
(a)	Teacher is bilingual.
(b)	Teacher uses both L1 and L2 for instruction.
(c)	L1 language arts skills are developed first, before introduction of L2 language arts.
(d)	L1 is use substantially for instruction. At least 50% of the total instructional time (approximately three hours per day) is in L1.
(e)	Use of L1 and L2 is differentiated by teaching staff (e.g., Teacher A only uses L1, and Teacher B only uses L2).
(f)	Teachers using L2 have "native" or "near-native" L2 skills.
(g)	This is a late-exit program. That is, children are not mainstreamed into English-only programs until the end of the fifth or sixth grade.
(h)	There is a primary language arts component.
(i)	There is an L2 language arts component.
(j)	Cultural sensitivity is reflected in the program by the teacher, the instructional materials, and the children's tasks.
(k)	Math is taught in L1.

The survey efforts resulted in the selection and participation of four immersion strategy sites and three late-exit transitional programs in Year 1 (FY 84-85) of data collection. Four early-exit transitional bilingual programs were selected from the four districts implementing an immersion strategy program.

In an effort to expand the number of immersion strategy programs, in Year 2 (FY 85-86), a new site was brought into the study, expanding the total number of immersion strategy and early-exit transitional bilingual locations from four to five each. In addition, in Year 2 (FY 85-86), to confirm the achievement patterns observed in the first year of data collection, new cohorts of children were added to the study from some of the districts in each program model which participated in Year 1 of data collection.

As can be seen in Table 5, the geographic distribution of the sites provides a balance of representation of the major Spanish-speaking groups in the country (i.e., Mexican; Mexican-American;

ERIC

Chicano in the west and southwest; and Puerto Rican and Cuban in the east) for each instructional program.

Table 5

Study Sites by Geographic Area and Program

Site	State	Immersion Strategy	Early-Exit Transitional	Late-Exit Transitional
A	Texas	X	X	
B	California	X	X	
C	California	X	X	
D	Florida			X
E	New York			X
F	Texas	X	X	
G	California			X
H	New Jersey	X		
I	New Jersey		X	

Cohorts

In the first year of data collection, the study design (Ramirez et al., 1984) called for a cohort of kindergartners from each of the three instructional programs, a cohort of first-graders from the immersion strategy and early-exit TBE programs, and a cohort of third-graders from the late-exit TBE program. These cohorts were selected on the basis that the goals of immersion strategy and early-exit programs are for rapid transition to mainstream programs, usually within two to three years of entry into the program. Thus, beginning the study with kindergarten and first grade cohorts allows for documentation of these two instructional programs as well as for the possible follow-up of students after a year or two in mainstream classrooms. Beginning the study with kindergarten and third grade cohorts in late-exit programs allows for documentation of the program from kindergarten to sixth grade. This design allows the testing of the model's thesis that the facilitative effects of instruction in the LEP child's primary language become apparent after grade four. As noted above, to effect a partial replication of first-year results, new kindergarten and third grade cohorts were begun in Year 2 of the study in some districts. And to provide for a greater range of sites implementing an immersion strategy program, a new site was brought into the study with a kindergarten and first grade cohort. Tables 6, 7, 8 and 9 list the cohorts and grade levels for Years 1, 2, 3 and 4 of the study (1986-87). The additions in FY 1985-86 modified the original study design (Ramirez et al., 1984). Figure 1 illustrates the resultant study design. No additional cohorts were added in Years 3 or 4 of data collection.

Table 6

Grade Levels Included for Each District
Year 1 of Data Collection, FY 84-85

District	Program(s)	K	1	2	3
A	Immersion Early-Exit	Cohort 10	Cohort 11	—	—
B	Immersion Early-Exit	Cohort 10	Cohort 11	—	—
C	Immersion Early-Exit	Cohort 10	—	—	—
D	Late-Exit	Cohort 10	—	—	Cohort 13
E	Late-Exit	Cohort 10	—	—	Cohort 13
F	Immersion Early-Exit	Cohort 20*	Cohort 21	—	—
G	Late-Exit	Cohort 20	—	—	Cohort 23

Cohorts 10, 11 and 13 - Students entering Fall 1984, Year 1 of data collection

Cohorts 20, 21 and 23 - Students entering Spring 1985, Year 1 of data collection

— - No students at this grade level in Year 1 of data collection, FY 84-85

* - Early Exit only in this cohort in this district.

Note: All kindergarten students begin program in EE at district F, then are arbitrarily assigned to IS or EE in first grade.

Table 7

Grade Levels Included for Each District
Year 2 of Data Collection, FY 85-86

District	Program(s)	K	1	2	3	4
A	Immersion Early-Exit	—	Cohort 10	Cohort 11	—	
B	Immersion Early-Exit	Cohort 30	Cohort 10	Cohort 11	—	
C	Immersion Early-Exit	Cohort 30	Cohort 10	—	—	—
D	Late-Exit	—	Cohort 10	—	—	Cohort 13
E	Late-Exit	Cohort 30	Cohort 10	—	Cohort 33	Cohort 13
F	Immersion Early-Exit	—	—	Cohort 21	—	—
G	Late-Exit	—	Cohort 20	—	—	Cohort 23
H	Immersion	Cohorts 30,40	Cohorts 31,41	—	—	—
I	Early-Exit	Cohort 30	Cohort 31	—	—	—

Cohorts 10, 11 and 13 - Students entering Fall 1984, Year 1 of data collection

Cohorts 20, 21 and 23 - Students entering Spring 1985, Year 1 of data collection

Cohorts 30, 31 and 33 - Students entering Fall 1985, Year 2 of data collection

Cohorts 40 and 41 - Students entering Spring 1986, Year 2 of data collection

— - No students at this grade level in Year 2 of data collection, FY 85-86

ERIC

Table 8

Grade Levels Included for Each District
Year 3 of Data Collection, FY 86-87

District	Program(s)	1	2	3	4	5
A	Immersion Early-Exit	—	Cohort 10	Cohort 11	—	—
B	Immersion Early-Exit	Cohort 30	Cohort 10	Cohort 11	—	—
C	Immersion Early-Exit	Cohort 30	Cohort 10	—	—	—
D	Late-Exit	—	Cohort 10	—	—	Cohort 13
E	Late-Exit	Cohort 30	Cohort 10	—	Cohort 33	Cohort 13
F	Immersion Early-Exit	—	—	Cohort 21	—	—
G	Late-Exit	—	Cohort 20	—	—	Cohort 23
H	Immersion	Cohorts 30,40	Cohorts 31,41	—	—	—
I	Early-Exit	Cohort 30	Cohort 31	—	—	—

Cohorts 10, 11 and 13 - Students entering Fall 1984, Year 1 of data collection
Cohorts 20, 21 and 23 - Students entering Spring 1985, Year 1 of data collection
Cohorts 30, 31 and 33 - Students entering Fall 1985, Year 2 of data collection
Cohorts 40 and 41 - Students entering Spring 1986, Year 2 of data collection
— - No students at this grade level in Year 3 of data collection, FY 86-87

ERIC

Table 9

Grade Levels Included for Each District
Year 4 of Data Collection, FY 87-88

District	Program(s)	2	3	4	5	6
A	Immersion Early-Exit	—	Cohort 10	Cohort 11	—	—
B	Immersion Early-Exit	Cohort 30	Cohort 10	Cohort 11	—	—
C	Immersion Early-Exit	Cohort 30	Cohort 10	—	—	—
D	Late-Exit	—	Cohort 10	—	—	Cohort 13
E	Late-Exit	Cohort 30	Cohort 10	—	Cohort 33	Cohort 13
F	Immersion Early-Exit	—	—	Cohort 21	—	—
G	Late-Exit	—	Cohort 20	—	—	Cohort 23
H	Immersion	Cohorts 30,40	Cohorts 31,41	—	—	—
I	Early-Exit	Cohort 30	Cohort 31	—	—	—

Cohorts 10, 11 and 13 - Students entering Fall 1984, Year 1 of data collection
Cohorts 20, 21 and 23 - Students entering Spring 1985, Year 1 of data collection
Cohorts 30, 31 and 33 - Students entering Fall 1985, Year 2 of data collection
Cohorts 40 and 41 - Students entering Spring 1986, Year 2 of data collection
— - No students at this grade level in Year 4 of data collection, FY 87-88

Figure 1

Immersion Strategy (IS), Early-Exit (EE), and Late-Exit (LE) Programs

By Grade, Cohort* and Year Entered Study

Year of Data Collection	Grade K	1	2	3	4	5	6
	Cohorts	Cohorts	Cohorts	Cohorts	Cohorts	Cohorts	Cohorts
1. 1984-85	10,20	11,21		13,23			
2. 1985-86	30,40	10,20,31,41	11,21	33	13,23		
3. 1986-87		30,40	10,20,31,41	11,21	33	13,23	
4. 1987-88			30,40	10,20,31,41	11,21	33	13,23
Cohort-Program:Districts			30-IS:B,C,H 40-IS:H 30-EE:B,C,I 30-LE:E	10-IS:A,B,C 20-IS:C 10-EE:A,B,C 20-EE:F** 10-LE:D,E 20-LE:G	11-IS:A,B 21-IS:F 11-EE:A,B 21-EE:F	33-LE:E	13-LE:D,E 23-LE:G
Year Entered Study			1985-86	1984-85	1984-85	1985-86	1984-85
Cohort-Program:Districts				31-IS:H 41-IS:H 31-EE:I			
Year Entered Study				1985-86			

* First digit of cohort number refers to study entry period (1: fall, 1984; 2: spring, 1985; 3: fall, 1985; and 4: spring, 1986). Second digit of cohort number refers to study entry grade (0: kindergarten, 1: first grade, and 3: third grade). For example, cohort 23 = student entering study in spring 1985 as a third grader.

** All students in cohort 20-EE:F exited study at end of kindergarten (due to loss of funds for program in district F).

NOTE: All kindergarten students in district F begin program in EE, then are arbitrarily assigned to IS or EE in first grade.

Target Schools and Teachers

Schools and teachers were selected for participation in the study using the following procedures (see Table 10).

Immersion Strategy Program. Except in one district, all schools and all kindergarten and/or first grade teachers within those schools implementing the immersion strategy program were selected to participate in the study. In one district, all schools implementing the immersion strategy program participated in the study and kindergarten and/or first grade teachers were selected arbitrarily by the district's project director for participation in the study. Once selected, all teachers voluntarily participated in the study.

Early-Exit Program. In the district with only an early-exit program and those with both immersion strategy and early-exit programs, schools were selected arbitrarily by the district's project director for participation in the study, with the following two exceptions. In one district, all schools implementing the early-exit program participated in the study. Secondly, all schools with both immersion strategy and early-exit programs participated in the study. Kindergarten and first grade teachers in schools implementing only the early-exit program as well as kindergarten and first grade early-exit teachers in schools implementing both programs were selected arbitrarily by the district's project director for participation in the study. However, in each of two districts, no early-exit teachers were selected in one school with both programs. Once selected, all teachers voluntarily participated in the study.

Late-Exit Program. In one district, kindergarten and third grade teachers were selected arbitrarily for participation in the study by the project director in the only school implementing the late-exit program. In the remaining two districts implementing the late-exit program, schools as well as kindergarten and third grade teachers within those schools were selected arbitrarily by the district's

ERIC

project director for participation in the study. Once selected, all teachers voluntarily participated in the study.

Table 10

Selection of Target Schools and Teachers

District	School's Program Implementation	Selection of Schools	Selection of Teachers
A	Immersion	All	All
	Early-Exit	Arbitrary	Arbitrary
	Both Immersion & Early-Exit	N/A	N/A
B	Immersion	N/A	N/A
	Early-Exit	Arbitrary	Arbitrary
	Both Immersion & Early-Exit	All	All Immersion/ Arbitrary Early-Exit*
C	Immersion	All	All
	Early-Exit	Arbitrary	Arbitrary
	Both Immersion & Early-Exit	All	All Immersion/ Arbitrary Early-Exit*
D	Late-Exit	All	Arbitrary
E	Late-Exit	Arbitrary	Arbitrary
F	Immersion	N/A	N/A
	Early-Exit	All	Arbitrary
	Both Immersion & Early-Exit	All	All Immersion/ Arbitrary Early-Exit*
G	Late-Exit	Arbitrary	Arbitrary
H	Immersion	All	Arbitrary
I	Early-Exit	Arbitrary	Arbitrary

* At one school with both immersion strategy and early-exit programs, no early-exit teachers were selected.

Target Students

In all programs, students were assigned to study classrooms in accordance with the district's regular procedures for assigning

ERIC

limited-English-proficient students to classrooms. In districts with both immersion strategy and early-exit programs, students were arbitrarily assigned to program classrooms; no student attributes were used to assign students to the two programs beyond the common criteria used to identify limited-English-proficient students: (1) home language; (2) oral English language proficiency; and (3) oral Spanish language proficiency. That is, no special preselection criteria were used to assign students to study classrooms. Study classrooms typically included students with a range of English language skills. While limited-English and fluent-English proficient students were in the majority native English speakers were also evident.

From the students in the study classrooms, potential target students were selected from only those who met the selection criteria. The selection criteria reflect the assumption that only those students who have been in a district's program without interruption would provide the clearest assessment of the effectiveness of the instructional program. To meet the selection criteria, students must:

- o have Spanish as a primary language;
- o have been classified as limited-English-proficient (LEP) upon entry into kindergarten (although students in grades one and three may be reclassified as fluent-English-proficient (FEP));
- o have been enrolled in the study district and placed in the program under study since November 1 of their kindergarten year;
- o have not been absent more than 40 school days in any one given year.

Tables 11, 12, 13 and 14 display the total number of districts, schools, classrooms, and students by program participating in the study through Year 4 of data collection.

ERIC

Table 11

Number of Study Students
Year 1 - 1984-85

	IMMERSION				EARLY-EXIT				LATE-EXIT				TOTAL			
	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.
K	3	7	16	270	4	12	26	448	3	12	15	252	7	29	57	970
1	3	4	16	182	3	8	17	226					3	9	33	408
2																
3									3	13	21	193	3	13	21	193
4																
5																
6																
TOTAL**	4	8	32	452	4	13	43	674	3	13	36	445	7	30	111	1571

* Number of schools includes all schools with target study students in program and mainstream program classrooms. Number of classrooms includes only program classrooms in which target study students are enrolled.

** The sum of the number of districts or schools at each grade level may not equal the total number of districts or schools because more than one grade level may be included in the same district or school.

Table 12

Number of Study Students
Year 2 - 1985-86

	IMMERSION				EARLY-EXIT				LATE-EXIT				TOTAL			
	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.
K	3	12	13	194	3	7	12	191	1	6	7	141	5	22	32	526
1	4	16	25	307	4	13	27	299	3	12	24	203	8	39	76	809
2	3	4	11	111	3	7	17	119					3	8	28	230
3									1	5	7	78	1	5	7	78
4									3	13	18	137	3	13	18	137
5																
6																
TOTAL**	5	20	49	612	5	16	56	609	3	14	56	559	9	46	161	1780

* Number of schools includes all schools with target study students in program and mainstream program classrooms. Number of classrooms includes only program classrooms in which target study students are enrolled.

** The sum of the number of districts or schools at each grade level may not equal the total number of districts or school because more than one grade level may be included in the same district or school.

Table 13

Number of Study Students
Year 3 - 1986-87

	IMMERSION				EARLY-EXIT				LATE-EXIT				TOTAL			
	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.
K																
1	4	13	17	172	3	8	15	162	1	6	7	108	6	24	39	442
2	4	17	20	226	4	13	24	197	3	12	19	159	8	40	63	582
3	3	4	7	81	3	7	14	97					3	8	21	178
4									1	5	7	72	1	5	7	72
5									3	13	16	99	3	13	16	99
6																
TOTAL**	5	20	44	479	5	17	53	456	3	14	49	438	9	47	146	1373

* Number of schools includes all schools with target study students in program and mainstream program classrooms. Number of classrooms includes only program classrooms in which target study students are enrolled.

** The sum of the number of districts or schools at each grade level may not equal the total number of districts or school because more than one grade level may be included in the same district or school.

ERIC

Table 14

Number of Study Students
Year 4 - 1987-88

	IMMERSION				EARLY-EXIT				LATE-EXIT				TOTAL			
	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.	# of Dist.	# of Sch.*	# of Cls.*	# of Std.
K																
1																
2	4	17	19	151	3	8	13	112	1	6	8	75	6	28	40	338
3	4	19	17	158	4	13	21	151	3	12	19	136	8	42	57	445
4	3	4	5	59	3	7	12	86					3	8	17	145
5									1	5	7	51	1	5	7	51
6									3	11	15	75	3	11	15	75
TOTAL**	5	20	41	368	5	16	46	349	3	14	49	337	9	46	136	1054

* Number of schools includes all schools with target study students in program and mainstream program classrooms. Number of classrooms includes only program classrooms in which target study students are enrolled.

** The sum of the number of districts or schools at each grade level may not equal the total number of districts or school because more than one grade level may be included in the same district or school.

ERIC

Data Collection Instruments and Procedures

As noted in Figure 1, some target students entered the study in FY 1984-85 and others in FY 1985-86. Consequently, for some students there are data over a four-year period and for others over a three-year period. Figures 2 and 3 list the data collection instruments, by grade level and year of the study for cohorts entering in FY 1984-85 and FY 1985-86, respectively.

This chapter summarizes the procedures for data collection in all four years of the study. For each target child included in the sample, a variety of data about the child's academic performance and skills, family background, teacher, classroom, school, and district were collected. The following lists the data collected for each child:

YEAR 1 DATA COLLECTION

Child-Level Data

- oral English language proficiency	IPT, Student Oral Language Observation Measure, Student Data Sheets
- English language arts	TOBE (kindergarten), CTBS (grades 1 and 3), Student Data Sheets
- English reading	TOBE (kindergarten), CTBS (grades 1 and 3), Student Data Sheets
- math assessed in English	TOBE (kindergarten), CTBS (grades 1 and 3), Student Data Sheets
- oral Spanish language proficiency	IPT, Student Data Sheets
- Spanish language arts	TOBE (kindergarten), CTBS (grades 1 and 3), Student Data Sheets

ERIC

Figure 2

Data Collection Plan
Cohorts Entering Study 1984-85

	YEAR 1 DATA COLLECTION (1984-85)						YEAR 2 DATA COLLECTION (1985-86)						YEAR 3 DATA COLLECTION (1986-87)						YEAR 4 DATA COLLECTION (1987-88)					
	Fall			Spring			Fall			Spring			Fall			Spring			Fall			Spring		
Testing:	K	1	3	K	1	3	1	2	4	1	2	4	2	3	5	2	3	5	3	4	6	3	4	6
TOBE-2 English				X																				
TOBE-2 Spanish	X			X																				
CTBS-S		X	X		X	X				X	X	X				X	X	X				X	X	X
CTBS-Espanol		X	X		X	X																X	X	X
IPT English	X	X	X	X	X	X				X	X	X				X	X	X				X	X	X
IPT Spanish	X	X	X	X	X	X																X	X	X
SOLOM	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
WISC-R			X											X					X					
Interviews:																								
Teacher				X	X	X				X	X	X				X	X	X				X	X	X
Aide				X	X	X				X	X	X				X	X	X				X	X	X
Parent				X		X											X							
Student Data				X	X	X				X	X	X				X	X	X				X	X	X
Classroom Observations:																								
Teacher Focus	X	X	X																					
Engaged Academic Time	X	X	X	X	X	X																		
Classroom Status Measure	X	X	X	X	X	X																		
Student Focus	X	X	X	X	X	X																		
Language Observation Measure										X	X	X				X	X	X				X	X	X
Classroom Engaged Academic Time Measure										X	X	X				X	X	X				X	X	X

Figure 3

Data Collection Plan
Cohorts Entering Study 1985-86

	YEAR 1 DATA COLLECTION (1984-85)	YEAR 2 DATA COLLECTION (1985-86)						YEAR 3 DATA COLLECTION (1986-87)						YEAR 4 DATA COLLECTION (1987-88)					
	Fall / Spring	Fall			Spring			Fall			Spring			Fall			Spring		
Testing:		K	1	3	K	1	3	1	2	4	1	2	4	2	3	5	2	3	5
TOBE-2 English	Cohorts entering study 1985-86 did not begin until Year 2 of Data Collection				X														
TOBE-2 Spanish		X			X														
CTBS-S			X	X		X	X				X	X	X				X	X	X
CTBS-Espanol			X	X		X	X										X	X	X
IPT English		X	X	X	X	X	X				X	X	X				X	X	X
IPT Spanish		X	X	X	X	X	X										X	X	X
SOLOM		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
WISC-R				X										X	X				
Raven's Coloured Progressive Matrices																	X	X	
Interviews:																			
Teacher					X	X	X				X	X	X				X	X	X
Aide					X	X	X				X	X	X				X	X	X
Parent					X		X											X	
Student Data					X	X	X				X	X	X				X	X	X
Classroom Observations:																			
Language Observation Measure					X	X	X				X	X	X				X	X	X
Classroom Engaged Academic Time Measure					X	X	X				X	X	X				X	X	X

- Spanish reading	TOBE (kindergarten), CTBS (grades 1 and 3), Student Data Sheets
- math assessed in Spanish	TOBE (kindergarten), CTBS (grades 1 and 3), Student Data Sheets
- class schedule	Teacher Interview, Student Data Sheets
- special needs	Teacher Interview, Student Data Sheets
- general ability	WISC-R/EIWN-R (grade 3)
Home Background	
- income	Parent Interview (kindergarten and grade 3)
- parent education, employment	Parent Interview (kindergarten and grade 3)
- home/community language usage	Parent Interview (kindergarten and grade 3)
- parent participation	Parent Interview (kindergarten and grade 3)
- parent attitudes toward development of Spanish and acquisition of English language skills	Parent Interview (kindergarten and grade 3)
- length of time in the U.S.	Parent Interview (kindergarten and grade 3)
Teacher	
- training	Teacher Interview
- experience	Teacher Interview
- English, Spanish usage in class	Teacher Interview, Teacher Focus, Student Focus, Engaged Academic Time Measure, Classroom Status Measure

	attitudes toward developing Spanish and English language skills	Teacher Interview
<u>Classroom Aide</u>		Classroom Aide Interview
<u>Classroom</u>		
-	proportion of students by language classification	Teacher Interview
-	limited-English-proficient, fluent-English-proficient, or English only	Teacher Interview
-	instructional materials	Teacher Interview
-	teacher/child interactions	Teacher Focus
-	engaged academic time	Engaged Academic Time Measure
-	student groups and activities	Classroom Status Measure
<u>School/District</u>		
-	proportion of language-minority schools	Project Administrator Interview, School Site Fact Sheet, School District Services Questionnaire
-	socioeconomic status of [illegible]ent body	Project Administrator Interview, School Site Fact Sheet, School District Services Questionnaire
-	availability of specialized programs and/or services	Project Administrator Interview, Site Administrator Interview, School Site Fact Sheet, School District Services Questionnaire

- community characteristics	Project Administrator Interview, School Site Fact Sheet, School District Services Questionnaire

YEAR 2 DATA COLLECTION

Child-Level Data

- oral English language proficiency	IPT, Student Oral Language Observation Measure, Student Data Sheets
- English language arts	TOBE (kindergarten), CTBS (grades 1, 2, 3 and 4), Student Data Sheets
- English reading	TOBE (kindergarten), CTBS (grades 1, 2, 3 and 4), Student Data Sheets
- math assessed in English	TOBE (kindergarten), CTBS (grades 1, 2, 3 and 4), Student Data Sheets
- oral Spanish language proficiency	IPT (students entering study in Year 2), Student Data Sheets
- Spanish language arts	TOBE (kindergarten), CTBS (grades 1 and 3 for students entering study in Year 2), Student Data Sheets
- Spanish reading	TOBE (kindergarten), CTBS (grades 1 and 3 for students entering study in Year 2), Student Data Sheets
- math assessed in Spanish	TOBE (kindergarten) CTBS (grades 1 and 3 for students entering study in Year 2), Student Data Sheets
- class schedule	Teacher Interview, Student Data Sheets

- special needs	Teacher Interview, Student Data Sheets
- general ability	WISC-R/EIWN-R (grade 3)
Home Background	
- income	Parent Interview (kindergarten and grade 3)
- parent education, employment	Parent Interview (kindergarten and grade 3)
- home/community language usage	Parent Interview (kindergarten and grade 3)
- parent participation	Parent Interview (kindergarten and grade 3)
- parent attitudes toward development of Spanish and acquisition of English language skills	Parent Interview (kindergarten and grade 3)
- length of time in the U.S.	Parent Interview (kindergarten and grade 3)
Teacher	
- training	Teacher Interview
- experience	Teacher Interview
- English, Spanish usage in class	Teacher Interview, Language Observation Measure, Classroom Engaged Academic Time Measure
- attitudes toward developing Spanish and English language skills	Teacher Interview

- oral English and Spanish language proficiency	Teacher Oral English Language Proficiency Interview Protocol, Teacher Oral Spanish Language Proficiency Interview Protocol (all teachers participating in study in years 1 and 2)
<u>Classroom Aide</u>	Classroom Aide Interview
<u>Classroom</u>	
- proportion of students by language classification	Teacher Interview
- limited-English-proficient, fluent-English-proficient, or English only	Teacher Interview
- instructional materials	Teacher Interview
- teacher/child interactions	Language Observation Measure
- engaged academic time	Classroom Engaged Academic Time Measure
- student groups and activities	Classroom Engaged Academic Time Measure
<u>School/District</u>	
- proportion of language-minority schools	Project Administrator Interview, School Site Fact Sheet, School District Services Questionnaire (for schools entering study in year 2)
- socioeconomic status of student body	Project Administrator Interview, School Site Fact Sheet, School District Services Questionnaire (for schools entering study in year 2)

- availability of specialized programs and/or services	Project Administrator Interview, Site Administrator Interview, School Site Fact Sheet, School District Services Questionnaire (for schools entering study in year 2)
- community characteristics	Project Administrator Interview, School Site Fact Sheet, School District Services Questionnaire (for schools entering study in year 2)

YEAR 3 DATA COLLECTION

Child-Level Data

- oral English language proficiency	IPT, Student Oral Language Observation Measure, Student Data Sheets
- English language arts	CTBS (grades 1, 2, 3, 4 and 5), Student Data Sheets
- English reading	CTBS (grades 1, 2, 3, 4 and 5), Student Data Sheets
- math assessed in English	CTBS (grades 1, 2, 3, 4 and 5), Student Data Sheets
- oral Spanish language proficiency	Student Data Sheets
- Spanish language arts	Student Data Sheets
- Spanish reading	Student Data Sheets
- math assessed in Spanish	Student Data Sheets
- class schedule	Teacher Interview, Student Data Sheets
- special needs	Teacher Interview, Student Data Sheets

- general ability	WISC-R/EIWN-R (grade 3)
Home Background	
- income	Parent Interview (grade 3)
- parent education, employment	Parent Interview (grade 3)
- home/community language usage	Parent Interview (grade 3)
- parent participation	Parent Interview (grade 3)
- parent attitudes toward development of Spanish and acquisition of English language skills	Parent Interview (grade 3)
- length of time in the U.S.	Parent Interview (grade 3)
Teacher	
- training	Teacher Interview
- experience	Teacher Interview
- English, Spanish usage in class	Teacher Interview, Language Observation Measure, Classroom Engaged Academic Time Measure
- attitudes toward developing Spanish and English language skills	Teacher Interview
- oral English and Spanish language proficiency	Teacher Oral English Language Proficiency Interview Protocol, Teacher Oral Spanish Language Proficiency Interview Protocol

Classroom

- proportion of students by language classification	Teacher Interview
- limited-English-proficient, fluent-English-proficient, or English only	Teacher Interview
- instructional materials	Teacher Interview
- teacher/child interactions	Language Observation Measure
- engaged academic time	Classroom Engaged Academic Time Measure
- student groups and activities	Classroom Engaged Academic Time Measure

YEAR 4 DATA COLLECTION

Child-Level Data

- oral English language proficiency	IPT, Student Oral Language Observation Measure, Student Data Sheets
- English language arts	CTBS (grades 2, 3, 4, 5 and 6), Student Data Sheets
- English reading	CTBS (grades 2, 3, 4, 5 and 6), Student Data Sheets
- math assessed in English	CTBS (grades 2, 3, 4, 5 and 6), Student Data Sheets
- oral Spanish language proficiency	IPT, Student Data Sheets
- Spanish language arts	CTBS (grades 2, 3, 4, 5 and 6), Student Data Sheets

- Spanish reading	CTBS (grades 2, 3, 4, 5 and 6), Student Data Sheets
- math assessed in Spanish	CTBS (grades 2, 3, 4, 5 and 6), Student Data Sheets
- class schedule	Teacher Interview, Student Data Sheets
- special needs	Teacher Interview, Student Data Sheets
- general ability	WISC-R/EIWN-R/Raven's CPM (grades 2 and 3)
<u>Home Background</u>	
- income	Parent Interview (grade 3)
- parent education, employment	Parent Interview (grade 3)
- home/community language usage	Parent Interview (grade 3)
- parent participation	Parent Interview (grade 3)
- parent attitudes toward development of Spanish and acquisition of English language skills	Parent Interview (grade 3)
- length of time in the U.S.	Parent Interview (grade 3)
<u>Teacher</u>	
- training	Teacher Interview
- experience	Teacher Interview
- English, Spanish usage in class	Teacher Interview, Language Observation Measure, Classroom Engaged Academic Time Measure

- attitudes toward developing Spanish and English language skills	Teacher Interview
- oral English and Spanish language proficiency	Teacher Oral English Language Proficiency Interview Protocol, Teacher Oral Spanish Language Proficiency Interview Protocol

Classroom

- proportion of students by language classification	Teacher Interview
- limited-English-proficient, fluent-English-proficient, or English only	Teacher Interview
- instructional materials	Teacher Interview
- teacher/child interactions	Language Observation Measure
- engaged academic time	Classroom Engaged Academic Time Measure
- student groups and activities	Classroom Engaged Academic Time Measure

Data Collection

Tests and Interviews

Once the sample selection was completed, a site coordinator was hired for each site. The site coordinator's primary responsibility was to schedule and supervise the achievement, language and general ability assessment testing, parent, teacher, and classroom aide interviews and completion of the student data sheets. To facilitate data collection, site coordinators were trained by senior study staff. Site coordinators then hired and trained local testers, proctors, and parent interviewers. Bilingual personnel were hired to ensure that only those

ERIC

proficient in English administered the English versions of tests and those proficient in Spanish administered the Spanish versions. Similarly, members of the community who were proficient in Spanish contacted parents and conducted the parent interviews. Given the time interval between fall and spring testing, senior Aguirre/SRA staff conducted on-site training of all test administrators and proctors prior to fall and spring testing to review all testing protocols and to ensure that test administration procedures were followed. Senior project staff completed all interviews of principals, project directors, and additional district personnel as needed.

Instrumentation used during the testing periods included standardized achievement and general ability tests and language proficiency tests in both English and Spanish.

The Tests of Basic Experiences (TOBE-2) was administered to kindergarten students in their first year in the study in their dominant language in the fall and in both English and Spanish in the spring. The Comprehensive Test of Basic Skills was administered to all students except kindergartners in their first year in the study in both English (CTBS-S) and Spanish (CTBS-Espanol) in the fall and spring. The CTBS-S was administered to all continuing target students in the spring of each successive year. The CTBS-Espanol was administered to all continuing target students in the spring of the fourth year of the study.

Standardized testing protocols were followed for TOBE and CTBS test administration. Directions regarding test preparation, test administrator/proctor to student ratio, testing environment, test instructions, and test section time limits were all read, reviewed and practiced by all test examiners during one-day training sessions.

The IDEA Oral Language Proficiency Test (IPT) English was administered to students in all grades to assess English oral language proficiency at fall and spring testings of their first year in the

ERIC

study and in spring of each successive year. The IPT Spanish was administered to assess Spanish oral language proficiency at fall and spring testings of their first year in the study and in spring of the fourth year of the study. The test administration protocol for the IPT was followed with one deviation. All Aguirre/SRA study students began the test with item number one. The examiner's manual directs the test examiner to begin the test at varying start levels for later grade students (i.e., $\geq$ grade 2). Aguirre/SRA's assumption was that all students would not necessarily be able to answer all items at beginning levels accurately.

The Student Oral Language Observation Matrix (SOLOM), a measure of students' oral English language proficiency, was also completed for each student by his or her classroom teacher. Teacher training in use of the SOLOM was conducted by senior project staff on-site at each school district.

The WISC-R/EIWN-R (Wechsler Intelligence Scale for Children-Revised/Escala de Inteligencia Wechsler para Ninos-Revisada) was administered to each target student participating in the study, usually in the course of his or her third grade year. Standardized testing protocols were followed for test administration and assessment of general ability. Bilingual (Spanish/English) certified clinical and school psychologists and advanced graduate students supervised by a certified school psychologist administered all test protocols. The Raven's Coloured Progressive Matrices was also administered to some students in one school district to assess general ability. All test administrators were trained by a certified school psychologist, supervised by the school district director of testing and measurement, and followed standardized testing protocols for test administration.

Site coordinators conducted all teacher and classroom aide interviews and supervised completion of all student data sheets. Aguirre/SRA staff provided training in procedures for completing these interviews. Training was also provided on-site at each district, to all bilingual interviewers assigned the task of completing parent interviews. Directions for scheduling and arranging the interviews, as well as techniques to elicit responses to items, were reviewed and practiced. All parent interview training was conducted bilingually.

Oral English and Spanish language interviews were conducted with all target classroom teachers participating in the study in FY 1984-85, FY 1985-86, FY 1986-87 and FY 1987-88. The objective was to describe the oral English and Spanish skills of teachers in the three instructional programs: immersion strategy, early-exit and late-exit bilingual programs. The interview protocol was adapted from a procedure developed by the Defense Language Institute/Language School. Teachers' proficiency in the comprehension and production of English and Spanish was assessed in the following areas: pronunciation, vocabulary, accuracy, contextual structure, fluency and participation in role-play situations. Interviews were conducted as informal conversations with the interviewer guiding the use of language at various proficiency levels. All interviewers were trained by Pelavin Associates staff under the direction and supervision of Aguirre International/SRA Technologies senior staff. The interview protocol, as well as elicitation and scoring techniques, were reviewed and practiced. All interviews were tape-recorded, and interrater reliability assessed.

All completed tests and interviews were sent to Aguirre/SRA for a final check for completeness and accuracy. Preparation of data for computer input was supervised by senior project staff.

Classroom Observations

Teacher Focus. The Teacher Focus observation instrument documents how teachers and students use language in the classroom. Those behaviors which research on first and second language development identify as facilitating language development were coded. Also coded were behaviors deemed important in research on teaching effectiveness. These data were collected in two steps. Step 1 consisted of tape recording classroom language as used by the teacher and students. Step 2 consisted of the transcription and coding of these tapes.

During Step 1, a wireless microphone was attached to the teacher and two stationary microphones were positioned on opposite ends of the classroom. These microphones permitted taping of both teachers and the student(s) with whom they were speaking. In each classroom, only a half day's instruction was taped — either in the morning or in the afternoon. Teachers were requested to schedule observations such that one period each of language arts, reading, math, and another content area were observed.

Completion of Step 2, transcription and coding of tapes, was conducted after the observations. All classroom observations took place during the fall of 1984 (study period 1). The following outlines the procedures for transcribing and coding of the classroom recordings.

A transcription was made of a random sample from each classroom recording. Each of the 70 Teacher Focus observations encompassed a maximum of 120 minutes on up to three sides of cassette tape (a set of two tapes), with a maximum of 45 minutes per side. From each classroom observation, a random 30 minutes of instruction was sampled according to the following procedure:

- the three tape sides were divided into 12, 10-minute portions in sequence, using the tape counter;

ERIC

- 29 different 30-minute sample designs were generated, each consisting of three randomly selected 10-minute portions from the total 120 minutes per classroom;
- the 70 classroom tape sets were grouped into the three programs (Program 1 = 21 tapes, Program 2 = 20 tapes, Program 3 = 29 tapes);
- one classroom tape set from each program was randomly assigned to each of 20 sample designs, and the remaining tape sets from Programs 1 and 3 were randomly assigned to the remaining designs;
- during transcription, any breaks or gaps in recording, i.e., blank tape that was to be sampled in a specific portion, was made up with alternate 10-minute portions from elsewhere on the tapes.

Following the selection of the tape sample, the 30-minute sequences from each tape were transcribed according to the detailed procedures outlined in Chaudron's (1985a), "Transcription Directions for SRA Project." A sample consisting of 70% of the transcribed tapes, representing each transcriber, was checked by a different transcriber to ensure accurate transcription, according to the procedures outlined in Chaudron (1985b), "Instructions for Verifying Transcripts." Hardcopy transcripts of the 30-minute samples were coded according to the instructions as outlined in Chaudron's (1985c), "Coding Procedures and Guidelines for SRA Project."

Four-page segments of the two transcripts were coded according to the guidelines by three coders who reached consensus on the codes. Three- to four-page segments of four additional transcripts were coded separately by two of these coders until 90% agreement was reached, and the remaining differences were used to establish modifications and clarifications of the guidelines, and a standard set of codes for the entire set of training transcripts. The first two transcripts were then discussed in group with all the coders to assure understanding of the guidelines.

Each coder then independently coded one segment of a transcript at a time, until 85% agreement with the standard codes was reached on a given training transcript. After reaching this criterion of reliability, and discussion of the remaining differences in coding, each coder proceeded to code, to the extent possible, the classroom transcripts that he or she had transcribed.

The codes on each transcript were then tallied according to the guidelines in Chaudron's (1985d), "Directions for Tallying Transcripts" with a running tally by page of each code. A second pass separated out the frequency of codes by language used.

Student Focus. The Student Focus Measure was used to describe the frequency with which students produce language in the classroom. Student utterances were coded in terms of with whom a student interacted (another student or a teacher/aide), who initiated the interaction, to whom the initiation was directed, whether the utterance was task-related, and the language used. In effect an utterance in this measure is defined as a speaker's turn in a turn-taking sequence.

Student Focus data were collected at the same moment that the Teacher Focus data were collected. Whereas the Teacher Focus data were coded from transcripts of live audiotapes of teacher/student interactions, the Student Focus measure was coded live. The objective was to describe the typical classroom day for target students in each instructional program. Classrooms were observed twice, once in the Fall, 1984 and then once in the Spring, 1985. Data were collected over approximately six hours of observation per classroom. All 70 classrooms in the study were observed in Fall, 1984. As the number of classrooms increased greatly during the year with the addition of new sites (from 70 to 111), 70 randomly sampled classrooms were observed in the Spring, 1985. Classrooms were sampled randomly from each program to ensure generalizability of study classrooms.

The Student Focus Measure was designed to complement the teacher-focused observations. The intent of the Teacher Focus data was to describe language in the classroom from the teacher's point of view.

This was done by coding all instances of teacher/student interaction. This allowed the description of how teachers use language in the classroom and their students' responses.

The student-focused observations allowed the description of how students use language in the classroom from the perspective of the student. The concern was to obtain a measure of the opportunities individual students have to produce language, with whom, about what, and in which language. Taken together, the teacher- and student-focused observations provide a more thorough understanding of how language is used in the classroom.

How were the Student Focus data collected? As noted, while one observer was recording classroom language of teachers and students by monitoring the teacher, a second observer was monitoring randomly-selected students. The Student Focus Measure was used in conjunction with the Engaged Academic Time Form and Classroom Status Observational Form. These instruments are described in detail below. After completing one round of the Engaged Academic Time and Classroom Status Measures, the Student Focus Measure was used for 30 minutes. The same six randomly-selected students who were observed for the Engaged Academic Time measures were observed again for the Student Focus Measure.

A separate Student Focus form was used for each of the six target students. Each target student was observed for a five-minute period. During this period, utterances made by the student and teachers were coded. The following was recorded for each utterance: (a) who produced the initiating utterance (target student or teacher); (b) what language was used; (c) was the student utterance on- or off-task; (d) was the student utterance spontaneous or a response; and (e) were teacher utterances directed to the target student, entire group in which target student was a member, or another student in the target student's group. When the five-minute period ended, the observer moved onto the second target student for another five-minute observation.

This procedure was repeated until all six target students were observed.

After completing this observation cycle, the observer took a five-minute break, at the end of which was started another round of Engaged Academic Time observations, with six new randomly-selected target students. Once the second Engaged Academic Time and Classroom Status observations were completed, another round of Student Focus observations were begun using the six new target students selected for the Engaged Academic Time Measure. Three rounds of Student Focus observations were completed per classroom per visit. Eighteen different students were thus observed at each visit. Classrooms were visited twice for half a day, once in the morning and once in the afternoon. In all, 70 classrooms were observed in the fall and spring, FY 84-85.

Engaged Academic Time and Classroom Status Measures. Two measures were used in conjunction with the Student Focus Observation Measurement, a measure of Engaged Academic Time and a measure of Classroom Status. The Engaged Academic Time Measure was used to record the instructional activities and task-engagement of target students. Its purpose was to enable judgments about the proportion of time each target student was "on-task." The Classroom Status Measure allowed the information collected on the Engaged Academic Time instrument to be set into the overall classroom context in which the task engagement was occurring. It was used to record the general instructional activity in the class that was concurrent with the behavior of the individual target students.

The observation process entailed an iterative cycle of observing six individual target students, randomly selected, and then the whole class. The cycle was repeated successively during a 20-minute observation period.

Each Engaged Academic Time observation of an individual target student consisted of recording eight data items:

- The group arrangement in which the target student was working;
- The type of activity the target student was engaged in;
- The type of teaching behavior being employed by the instructor toward the target student;
- Which instructor was displaying the teaching behavior;
- The content area of the target student activity;
- The language of the written instructional materials;
- The oral language being used by the target student, if any;
- The oral language being used by the instructor, if any;
- Whether or not the target student was on-task;

Each classroom status observation consisted of recording six data items for each individual student or group in the entire class:

- The number of students in the group;
- The group activity;
- The content of the activity;
- Which instructor was available to the group;
- The oral language being used by the group, if any;
- Whether target students were members of the group;

Language Observation Measure. The Language Observation Measure documents how teachers and students use language in the classroom. Coded are those behaviors which research on first and second language development identify as promoting language development. In Year 1 of the data collection, those language behaviors were gathered through tape recordings of classroom language as used by teachers and students. These recordings were subsequently transcribed and coded. As the use of tape recording is costly in time and money, the first year results of classroom observations were used to develop the Classroom Language Observation Measure to document the most salient of these language

ERIC

behaviors through live observation. This change permitted the expansion of the amount of information available per classroom as well as to accommodate the substantial increase in the number of study classrooms from Year 1 (111) to Year 2 of the study (160). The Classroom Language Observation Measure also was used to document language behaviors in project classrooms in Years 3 (N = 145) and 4 (N = 115) of the study.

The purpose of the Language Observation Measure is to enable observers to accurately document language use in the classroom. The Language Observation Measure helps observers to focus systematically on important aspects of language use across the instructional day. The result should provide a picture of the nature of classroom communication in programs serving limited-English-proficient students.

The Language Observation Measure is structured to describe the communication between teachers and students. The unit of analysis is time based. That is, on the minute, the observer looks up and following explicit rules, first identifies the beginning of an exchange. Secondly, the observer then codes certain language behaviors within a specific category system. The language behaviors are referred to in this discussion as utterances. As previously defined, an utterance is a speaker's turn in a turn-taking sequence (i.e., communicative exchange). This time-sampling approach to observing teacher language has been used in teacher effectiveness studies (Program on Teaching Effectiveness - Stanford University, 1976). The observer codes a number of characteristics about the teacher-student interaction: the observer codes who initiated the exchange; the language used; and the specific language behavior used to start the exchange. Then the respondent, the language used, and the specific responding language behavior are identified. In those instances where the initiator of the exchange reacts to the response, the specific follow-up language behavior is coded along with the specific language used. The exchange is described further by documentation of the following for each exchange:

ERIC

Function. The purpose or overall intent of the exchange;

Focus. Whether the exchange is concerned with how things are said (form) vs. with what is said (concept);

Activity. The sanctioned classroom activity providing the context for the exchange;

Realia. Whether visuals, objects, and/or physical gestures are used to support the language behaviors exhibited;

Content. The subject area that is being taught at the time the observation is made.

These last behaviors provide the contextual framework for understanding the specific language behaviors noted for each exchange (turn-taking sequence).

Classroom Engaged Academic Time Measure. The Classroom Engaged Academic Time Measure was used in conjunction with the Language Observation Measure to record the activities in the classes observed. While the Language Observation Measure involves the recording of the speaking behaviors of teacher and students, the Classroom Engaged Academic Time Measure allows these language behaviors to be set into the overall classroom context in which they are occurring. The Classroom Engaged Academic Time Measure also provides a measure of task engagement for each group of students within the classroom.

The Classroom Engaged Academic Time Measure is coded in the same cycle as the Language Observation Measure. That is, after completing a cycle of recording language behaviors using the Language Observation Measure, the Classroom Engaged Academic Time Measure is used to record a "snapshot" of the whole class before going on to the next cycle of frames of the Language Observation Measure. The Classroom Engaged Academic Time Measure allows the following information to be recorded:

- o The number of student groups;
- o The type of activity for each student group;

- The content of the activity for each student group;
- The language being used, if any, in each student group;
- The number of children wearing blue arm bands (LEP) who are engaged in each student group;
- The number of children wearing blue arm bands (LEP) who are not engaged in each student group;
- The number of children wearing red arm bands (FEP) who are engaged in each student group;
- The number of children wearing red arm bands (FEP) who are not engaged in each student group;
- The number of children wearing yellow arm bands (English only) who are engaged in each student group; and finally,
- The number of children wearing yellow arm bands (English only) who are not engaged in each student group.

A total of 160 classrooms in the second year, 145 in the third year, and 115 in the fourth year were observed both in the morning and afternoon. The goal was to obtain a picture of how language is used across the instructional day. Observations were not made during recess or lunch.

Only teachers of target classrooms, team teachers, or aides that were scheduled to work with target students on a daily basis were observed. The total amount of time each of these adults was observed was proportional to the amount of time each worked directly with target students.

While in each classroom, the observer coded four observation booklets per hour. Each booklet was divided into two parts. Part 1 was composed of the Classroom Language Observation Measure in which the observer coded teacher-student language use for ten minutes. At the end of this time, the observer proceeded to code classroom engaged academic time for five minutes in Part 2 of the observation booklet. Thus, each observation booklet represented a fifteen minute observation cycle. Four such cycles were completed for each hour of observation.

The total amount of observation time for each classroom averaged approximately 5.5 hours.

Bilingual observers were trained over a ten-day period which included discussion and field practice. Observer reliability was determined by comparing their coding of classroom transcripts in English and Spanish with those of expert coders. Observers were tested to determine the extent to which they correctly selected speech samples and coded specific language behaviors. Written transcripts were used, rather than audio and video tapes to ensure that observers were focusing on the same behaviors. Reliability checks were done at the end of training, and at the end of data collection. Only those observers with a reliability score of over 0.86 were selected as classroom observers. All estimates of observer reliability exceeded 0.86.

Generalizability

To what extent are study results applicable to all structured immersion strategy, early-exit, and late-exit transitional bilingual education programs?

At the time that this study was initiated, study sites represented the universe of immersion strategy programs, and all but one of the four late-exit programs in the United States as defined in this study. Given the diversity of Spanish-speaking groups (Mexican, Mexican-American/Chicano, Cuban, and Puerto Rican), the number of participating students (n = 2,352), the geographic distribution of study sites (rural, suburban, and urban), the large number of classrooms (n = 554), the comprehensiveness of the data collected (see Chapter III), and the consistency of the analytic results, one can be very confident that the study results adequately reflect the achievement of immersion strategy and late-exit students. Thus, generalizations for policy can be made on the basis of these numbers provided that the instructional programs as defined in this study represent acceptable instructional parameters for policy makers.

Early-exit programs were selected only from school districts that also had an immersion strategy program. Thus, the generalizability of the results of the early-exit program is best restricted to those early-exit programs having similar program characteristics to the early-exit programs in the study and that co-exist with immersion strategy programs. Moreover, since the time that this study was initiated, federal and state policy changes have allowed the implementation of English-only immersion strategy programs serving limited-English-proficient students. However, little or no data are available documenting the characteristics of these "newer" immersion strategy programs. Study results are generalizable to these newer programs to the extent that they resemble those immersion strategy programs included in this study (e.g., teachers are bilingual and have bilingual teaching credentials).

As a note, it is possible to use the data in this study to evaluate (i.e., model) how these newer or proposed new programs might affect the learning of Spanish-speaking limited-English-proficient students if the characteristics of these programs fall within the range of program characteristics observed in this study.

<u>To what extent are the results of this study applicable to all limited-English-proficient students and to all instructional programs designed to meet their needs</u>?

It is critical that the reader remember throughout this report that the goal, and, hence, the design of this study are to compare the relative effectiveness of three specific instructional programs serving Spanish-speaking limited-English-proficient elementary school students. This study is not designed to address the issue of the relative effectiveness of all instructional programs developed to serve limited-English-proficient students in the United States, nor is it designed to assess the overall effectiveness of each individual program. This question is addressed by another federal study (Development Associates, 1984). As such, the results of this study are only statistically generalizable to Spanish-speaking limited-English-proficient students having similar background characteristics as those students participating in this study, and who are in instructional programs having similar characteristics as those examined by this study. For example, study results cannot be generalized to limited-English-proficient students having a primary language other than Spanish (e.g., Hmong, Cantonese, Vietnamese, Portuguese, etc.), or to instructional programs provided to Spanish-speaking limited-English-proficient students in high school.

III. DATA ANALYSIS OVERVIEW

Analytic Approach

This chapter will discuss the analysis of the data collected during all four years of the study. The data collection instruments that were included in the analyses will be described. Also included will be a discussion of how data were summarized to the classroom level for two of the observation instruments. Several design and statistical issues which had an impact on the analysis will be addressed.

Data Collection Instruments Used for the Analyses

Language Observation Measure
Classroom Engaged Academic Time Measure
Teacher Focus
Classroom Data Base
Teacher Interview
Student Data Base
Student Data Sheet
Project Administrator
School District Services
Site Administrator (Principal)
School Site Fact Sheet
Parent Interview

WISC/Raven *
TOBE-2 *
CTBS-S (English) *

* No descriptive analyses

Data Collection Instruments Not Included in the Analyses

Classroom Status Measure
Student Focus
Engaged Academic Time Measure
Classroom Aide Interview
IPT English/Spanish
CTBS-Espanol (Spanish)

ERIC

The first three instruments were not included in the analyses as other data were more appropriate. These instruments were superceded by other instruments after the first year of the study. The aide interview was not included as these data were only available for a small number of classrooms in a few districts.

Summarizing Observational Instruments To The Classroom Level

There were two primary observation instruments used to gather information on the classrooms in the study. One of the instruments was based on observation of the teacher (Language Observation Measure), and the other instrument was based on observation of the classroom as a whole (Classroom Engaged Academic Time Measure). The data from each instrument were summarized at the classroom level, and these classroom summaries were analyzed by program and grade level.

The data for the Language Observation Measure were aggregated across students or teachers within each classroom. The classroom summary was produced by calculating for each classroom the percentage of observation "frames" that contained a given value for a variable. For example, if 200 frames were observed for a classroom and English was the indicated language in 150 of them, a "percentage English" for that classroom was calculated as 75%. Similarly, the "percentage Spanish," "percentage both," and "percentage non-language" was calculated for the classroom, with the percentages for the four possible language values (responses) necessarily adding to 100%. This process was then repeated for each of the variables on the observation instrument. The resulting classroom-level percentages were then compared across programs and grades.

The instrument based on observation of the classroom as a whole was summarized slightly differently. For the Classroom Engaged Academic Time Measure (CEATM), the classroom was divided into groups and information about all the groups present at a moment in time was recorded on a single page. In this way, a single page "snapshot" of

ERIC

the classroom could contain from one to twelve "frames," each frame corresponding to a group in the classroom.

If the CEATM data had been summarized in the same way as the Language Observation Measure — by calculating the percentage of frames for a classroom with a given value on a variable — undue weight would have been given to the frames with few students. For example, five pages each with only one large group of students (and, therefore, each with only one frame) would collectively count only as much as a single page with five smaller groups and, therefore, five frames. This was adjusted for by treating each page as an observation rather than treating each frame on a page as an observation. Therefore, for each page the percentage of frames with a given response to a variable (such as English on the language variable) was calculated, and then these percentages were averaged over all pages for a classroom. This produced the classroom summary that was analyzed across programs and grades.

Nominal Program Definitions

The primary study objective is to assess the relative effectiveness of three specific instructional programs. The programs in the study were selected based on criteria associated with each program model. For the program or programs at each site, the nominal program labels were determined by the Project Director after a personal interview/observation of the principal, teachers, parents and classrooms to determine the extent to which the specific program characteristics for a given model in this study were evident in the school program(s). The nominal program labels are an important treatment variable regardless of variation within programs.

Confounding Of District With Program

There are statistical problems which can result from aggregation of the data (the ecological fallacy). It is possible that by aggre-

ERIC

gating the data, a spurious yet statistically significant effect which is due to the influence of some unknown uncontrolled factor might be found. Since the program effect is of primary interest, aggregating to the district level might produce spurious effects and might obscure true program effects.

The small number of school districts in the U.S. having an immersion strategy program and the even smaller number having a late-exit program makes comparing the three programs difficult. A district either had both an immersion strategy and an early-exit program or had only a late-exit program. No district had all three programs. The original intention was to "partition out" the between-site (district) source of variance in the analyses; however, because districts are neither nested within programs nor completely crossed with programs, there is a confounding of district and program. As a result, there is no suitable way of getting independent estimates for program and district.

In order to assess the district effect, the final model for the immersion strategy and early-exit programs was tested for district-program interaction effects.

Confounding Of Grade Level With Program

Due to the design of the study, data were collected from kindergarten (K) through fourth grade students in the immersion strategy and early-exit programs, and from kindergarten through sixth grade students in the late-exit program. The immersion strategy and early-exit programs had cohorts who entered in kindergarten and first grade on the basis that one program goal is the rapid transition to mainstream programs. The late-exit program had cohorts who entered in kindergarten and third grade in order to document the program from kindergarten through sixth grade and to test the program thesis that the facilitative effects of instruction in the LEP child's primary language become apparent after grade four. While this design is effective for some purposes, it causes several analytic difficulties because of the number of empty cells in the design.

For the descriptive analyses, the discussion has focused on trends across grade levels and program differences for grades K-4. In most cases, the late-exit grades five and six followed the same trend (increasing, decreasing, or constant) as for late-exit grades K-4.

Unit Of Analysis

Because the primary analytic goal is the modeling of student achievement over time and because these data are collected at the individual student level, the most appropriate unit of analysis is the individual child. Although the programs are implemented at the classroom level and much of the data are collected at this level (or above), children do not remain in intact classes over time so it is impossible to use the classroom as the unit of analysis over four years. In addition, many of the most robust predictors of achievement in the educational research literature have been student-level variables such as socioeconomic status. In order to determine whether any of the teacher, classroom, school, and district variables predict achievement, these variables were linked to each child.

Descriptive Analyses

The goals of the descriptive analyses are: (a) to describe the treatments as perceived by parents, teachers, and school and district administrators, and as implemented in the classrooms; (b) to document that the treatments do, in fact, differ from one another in ways that reflect differences in their respective models; (c) to describe the students in the study in terms of home environment and teacher perceptions; and (d) to describe differences between students in the three programs on any variables which were considered for inclusion in subsequent analyses.

Weighting For Descriptive Analyses

Because the number of students, teachers, classrooms, schools, and administrators varied across districts, the proportion of data each district/program contributed varied across data collection instruments. For example, one district might have 10% of the student data sheets, 20% of the school site fact sheets, and 14% of the ICM observations. In order to represent more accurately the students, teachers, classrooms, schools, and administrators who participated in the study, all data were weighted proportional to the number of students in each district and program for all descriptive analyses. Appendix A describes the weighting algorithm used.

ERIC

IV. PROGRAM CHARACTERISTICS

Introduction

The purpose of this chapter is to provide additional information to the various questions of program implementation: Do the three programs under study in fact represent three distinct instructional programs? And if they do represent three different programs, do they represent the three programs of interest in this study--immersion strategy, early-exit, and late-exit transitional bilingual education programs? Also of interest is whether the programs differ on important variables unrelated to differences in the program models, but which might affect student outcome (e.g., whether the teachers in one program are more experienced than those in the other programs).

In this chapter information is presented on classroom language use, engaged academic time, teacher/classroom characteristics and administrative context.

Classroom Language

In years 2, 3 and 4 of data collection, classroom language observations were gathered through live coding of teacher-student language interactions. The questions addressed were: (1) Which language(s) do teachers and students use within the classroom? (2) What is the relationship of specific teacher language behaviors to specific student language behaviors? and (3) How do teachers use language for instruction in the classroom? Data for study year 1 which answers the question of which language(s) teachers and students use within the classroom are found in Appendix B. These data were collected in live audiorecording.

ERIC

Which language is used?

As defined in this study, immersion strategy, early-exit, and late-exit transitional bilingual programs are differentiated primarily on the basis of the language(s) used in the classroom and the amount each is used. All formal instruction in immersion strategy classrooms should be in English, with little or no use of Spanish. Both English and Spanish are used for instruction in early-exit and late-exit classrooms. However, according to the instructional model, Spanish should be used at least 40% of the time through grade 6 in late-exit classrooms. In contrast, teachers in early-exit programs dramatically taper the use of Spanish for instruction after kindergarten.

Does the language use pattern observed in the study's programs match the patterns described above?

Yes, for the immersion strategy and early-exit programs, but not quite for late-exit programs. Across all grades studied (i.e., K-4), immersion strategy teachers almost always use English (≥97.3%) and rarely use Spanish (≤1.5%) (see Table 15). While early-exit kindergarten and first grade teachers use English about two-thirds of the time (65.8% and 69.1%, respectively), the use of English increases rapidly to approximately three-fourths of the time in grades two (74.5%) and three (80.3%). By grade 4 English is used almost exclusively (97.3%), as it is among immersion strategy teachers. Conversely, early-exit teachers use Spanish less than one-third of the time in kindergarten (30.8%), quickly decreasing its use until it is seldom used in grade four (2.0%). These language use patterns for each program are consistent with the immersion strategy and early-exit models as defined in this study.

While late-exit teachers in the early grades, i.e., kindergarten to four, use English and Spanish in a manner consistent with the late-exit model as defined in this study, teachers in grades five and six did not. The data in Table 15 show how Spanish is used by late-exit teachers a large part of the time in kindergarten (87.8%), and its use slowly decreases to 42.6% in grade four. Contrary to the late-exit model, Spanish use continues to decrease to one-third in grade five (33.9%) and decreases even further in grade six (16.0%). This low level of Spanish use in the upper grades raises questions as to fidelity of treatment as defined (Spanish use ≥40%), as well as concern over the potential inhibiting effect upon primary language development. The grades in which use of Spanish conformed to the late-exit model (i.e., K-4), are those grades for which we also have data from immersion strategy and early-exit classrooms.

The consistency of observed language use patterns with the respective program models and the overlap of kindergarten through fourth grade levels allows assessment of the relative effectiveness of the three programs through these grade levels. The disruption in the observed language use pattern in grades five and six in the late-exit program, when compared to the model, might appear to limit the assessment of the program's effectiveness on achievement at these later grades for the late-exit model. However, the disruption reflects deviation from the late-exit model by only one of the three late-exit districts. Thus, there are sufficient data for late-exit fifth and sixth grade classrooms.

Table 16 describes the proportion of time (i.e., % of total discrete observations) that a teacher interacts with students in the classroom. Across the day, teachers produce non-verbal behaviors from 5.7% to 23.7% of the time (i.e., are not verbally interacting with students). In each program there is a slight trend among teachers to verbalize less as grade level increases. Notwithstanding, the patterns of English usage by program noted in Table 15 remain.

Table 15

Mean** Proportion of Teacher Utterances by Language, Program, and Grade

Grade		Program		
		Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 158.1	$\bar{X}$ = 191.5	$\bar{X}$ = 122.3
K	% English	98.5	65.8	9.3
	% Spanish	1.1	30.8	87.8
	% Mixed	0.4	3.4	3.0
		$\bar{X}$ = 207.3	$\bar{X}$ = 204.3	$\bar{X}$ = 178.8
1	% English	97.3	69.1	32.9
	% Spanish	1.5	28.7	65.2
	% Mixed	1.2	2.2	1.9
		$\bar{X}$ = 212.8	$\bar{X}$ = 211.0	$\bar{X}$ = 194.9
2	% English	98.2	74.5	30.3
	% Spanish	1.2	23.8	68.3
	% Mixed	0.6	1.7	1.4
		$\bar{X}$ = 177.1	$\bar{X}$ = 178.5	$\bar{X}$ = 175.0
3	% English	99.0	80.3	50.6
	% Spanish	0.4	16.8	45.4
	% Mixed	0.5	2.9	4.0
		$\bar{X}$ = 182.8	$\bar{X}$ = 197.3	$\bar{X}$ = 163.6
4	% English	99.8	97.3	55.3
	% Spanish	0.0	2.0	42.6
	% Mixed	0.2	0.7	2.1
				$\bar{X}$ = 179.6
5	% English	*	*	63.6
	% Spanish	*	*	33.9
	% Mixed	*	*	2.5
				$\bar{X}$ = 136.5
6	% English	*	*	80.3
	% Spanish	*	*	16.0
	% Mixed	*	*	3.7

** Mean teacher utterance is defined as the average number utterances made by teachers during the period of observation, approximately 5 hours.

Table 16

Mean** Proportion of Teacher Behaviors by Language, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 185.2	$\bar{X}$ = 215.2	$\bar{X}$ = 129.7
K	% English	84.5	58.0	9.0
	% Spanish	1.0	27.3	82.4
	% Mixed	0.3	3.1	2.9
	% Non-Verbal	14.2	11.6	5.7
		$\bar{X}$ = 225.1	$\bar{X}$ = 226.1	$\bar{X}$ = 194.1
1	% English	89.2	62.5	30.5
	% Spanish	1.3	25.6	59.9
	% Mixed	1.1	2.0	1.8
	% Non-Verbal	8.3	10.0	7.8
		$\bar{X}$ = 237.7	$\bar{X}$ = 238.5	$\bar{X}$ = 219.8
2	% English	87.2	66.1	26.8
	% Spanish	1.0	20.4	59.7
	% Mixed	0.5	1.4	1.3
	% Non-Verbal	11.3	12.1	12.3
		$\bar{X}$ = 212.1	$\bar{X}$ = 213.6	$\bar{X}$ = 200.4
3	% English	81.7	67.0	43.4
	% Spanish	0.4	13.6	40.3
	% Mixed	0.5	2.4	3.5
	% Non-Verbal	17.4	16.9	12.9
		$\bar{X}$ = 225.1	$\bar{X}$ = 237.5	$\bar{X}$ = 190.3
4	% English	80.2	80.9	48.0
	% Spanish	0.0	1.7	36.6
	% Mixed	0.2	0.6	1.8
	% Non-Verbal	19.6	16.9	13.6
				$\bar{X}$ = 207.6
5	% English	*	*	54.1
	% Spanish	*	*	29.3
	% Mixed	*	*	2.1
	% Non-Verbal	*	*	14.5

Table 16 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
				$\bar{X}$ = 178.2
6	% English	*	*	60.6
	% Spanish	*	*	12.7
	% Mixed	*	*	3.0
	% Non-Verbal	*	*	23.7

** Mean teacher behavior is defined as the average number behaviors made by teachers during the period of observation, approximately 5 hours.

<u>Does the student's use of English or Spanish differ between programs</u>?

Yes. Table 17 notes that student language use patterns are similar to those of their teachers. Immersion strategy students across all grades almost always use English (≥94.3%). Early-exit students in kindergarten and first grade use English approximately two-thirds of the time (65.6% and 64.5%, respectively), while those in second and third grade use it about three-fourths of the time (71.5% and 77.2%, respectively). By the fourth grade, early-exit students almost always use English (96.0%). In contrast, late-exit students in kindergarten use English sparingly (9.3%). First and second grade late-exit students use English about one-third of the time (32.5% and 30.8%, respectively). There is a big jump in the use of English in the third grade (52.0%). English usage continues to increase rapidly to about two-thirds in grades four (59.3%) and five (65.3%), until it is almost always used in grade six (83.3%). In sum, across grades and within program, students appear to mirror the language use patterns of their teachers.

ERIC

Table 17

Mean** Proportion of Student Utterances by Language, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 79.0	$\bar{X}$ = 87.3	$\bar{X}$ = 62.3
K	% English	94.3	65.6	9.3
	% Spanish	4.3	33.3	89.3
	% Mixed	1.4	1.1	1.5
		$\bar{X}$ = 96.7	$\bar{X}$ = 98.8	$\bar{X}$ = 82.0
1	% English	96.6	64.5	32.5
	% Spanish	2.6	34.0	66.8
	% Mixed	0.7	1.5	0.7
		$\bar{X}$ = 104.8	$\bar{X}$ = 96.6	$\bar{X}$ = 93.7
2	% English	97.2	71.5	30.8
	% Spanish	2.1	27.0	66.7
	% Mixed	0.7	1.4	2.4
		$\bar{X}$ = 83.3	$\bar{X}$ = 77.8	$\bar{X}$ = 82.9
3	% English	98.6	77.2	52.0
	% Spanish	0.5	20.3	45.2
	% Mixed	0.9	2.6	2.8
		$\bar{X}$ = 82.5	$\bar{X}$ = 90.6	$\bar{X}$ = 78.5
4	% English	98.0	96.0	59.3
	% Spanish	0.5	3.2	40.1
	% Mixed	1.5	0.8	0.6
				$\bar{X}$ = 90.2
5	% English	*	*	65.3
	% Spanish	*	*	32.2
	% Mixed	*	*	2.5
				$\bar{X}$ = 67.4
6	% English	*	*	83.3
	% Spanish	*	*	13.4
	% Mixed	*	*	3.3

** Mean is defined as the average frequency of student utterances recorded during the period of observation, on the average of 5 hours.

It is critical that language learners have the opportunity to produce language if they are to develop language. This opportunity to produce language is even more critical for second language learners who have to acquire a second language in a shorter time period and under more demanding contextual settings (i.e., classrooms) than first language learners (Cummins & Swain, 1986). Table 18 clearly documents, across grades within programs, that in almost half or more of teacher/student interactions, students do not produce any language (40.7% to 58.6%). Fr[illegible] a slightly different perspective, when the frequency of teacher and student utterances are compared, teachers do most of the talking, producing about two times as many utterances as do students ($\bar{X}$ = 122.3 to $\bar{X}$ = 212.8 vs. $\bar{X}$ = 62.3 to $\bar{X}$ = 104.8) (see Tables 15 and 17). Again, typically, three-fourths or more (72.0% to 81.9%) of student interactions with teachers are responses to teacher initiations (see Table 19). In sum, consistently across grades and program, second language learners are being provided with limited opportunities to produce language.

Table 18

Mean** Proportion of Student Behaviors [illegible] Language, Program, and Grade

Grade		Program: Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 152.1	$\bar{X}$ = 173.3	$\bar{X}$ = 107.0
K	% English	53.1	33.4	5.4
	% Spanish	2.6	14.7	53.1
	% Mixed	0.8	0.6	0.8
	% Non-Verbal	43.5	51.3	40.7
		$\bar{X}$ = 171.4	$\bar{X}$ = 184.5	$\bar{X}$ = 151.3
1	% English	54.4	35.3	17.8
	% Spanish	1.5	17.8	37.1
	% Mixed	0.5	0.8	0.4
	% Non-Verbal	43.7	46.1	44.7

Table 18 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 185.5	$\bar{X}$ = 189.0	$\bar{X}$ = 172.8
2	% English	54.5	37.3	16.5
	% Spanish	1.2	12.8	35.4
	% Mixed	0.4	0.7	1.4
	% Non-Verbal	43.9	49.2	46.7
		$\bar{X}$ = 174.4	$\bar{X}$ = 181.9	$\bar{X}$ = 163.0
3	% English	46.8	33.0	26.1
	% Spanish	0.3	8.6	24.9
	% Mixed	0.5	1.2	1.4
	% Non-Verbal	52.5	57.3	47.5
		$\bar{X}$ = 199.4	$\bar{X}$ = 204.0	$\bar{X}$ = 152.3
4	% English	40.5	42.4	31.7
	% Spanish	0.2	1.5	22.3
	% Mixed	0.7	0.4	0.3
	% Non-Verbal	58.6	55.6	45.7
				$\bar{X}$ = 166.7
5	% English	*	*	34.3
	% Spanish	*	*	17.6
	% Mixed	*	*	1.5
	% Non-Verbal	*	*	46.7
				$\bar{X}$ = 162.4
6	% English	*	*	33.9
	% Spanish	*	*	6.1
	% Mixed	*	*	1.7
	% Non-Verbal	*	*	58.3

** Mean is defined as the average frequency of student behaviors recorded during the period of observation, on the average of 5 hours.

What do teachers say?

Teacher utterances were coded into one of seven categories: (a) explaining--providing information; (b) questioning--asking students information; (c) commanding--directing students to do something; (d)

ERIC

modeling—demonstrating how something should be said; (e) feedback—informing students about their performance; (f) other—social comments such as, "How are you today," "What a nice picture," etc.; and (g) monitoring—teacher is not speaking at the time of observation, but is supervising students, such as during a test or seatwork.

Without considering the language used, no clear differences in patterns emerge between programs or grade levels, with two minor exceptions (see Table 20). The most common types of utterances are the same for all three programs: question, command, explanation, and feedback. Two exceptions are that while the use of modeling decreases, the use of explaining increases as grade level increases among immersion strategy classrooms. Additionally, teachers in each program tend to spend more of their time monitoring, i.e., greater use of passive instructional procedures, as grade level increases. These exceptions notwithstanding, the basic teaching behaviors appear to be equally present in about the same frequency across programs and grades.

Teaching behaviors are examined by language to determine if teachers use English and Spanish differently. Table 21 provides the distribution of teacher utterances when English is spoken. As before, across programs and grades the primary teaching behaviors are questioning, feedback, explaining and commanding. Only slight differences are noted across grades within programs. Across programs and grades, immersion strategy and early-exit teachers tend to explain more as grade level increases. No consistent pattern is noted across grades within program for late-exit teachers when they use English.

Program comparisons by grade show that when early-exit teachers use English, they tend to provide more explanations across grades when compared to late-exit teachers. While early-exit teachers also provide more explanations than do immersion strategy kindergarten and first grade teachers, no such program differences occur after grade one.

ERIC

Table 19

Mean Proportion of Student Utterances by Type, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 72.7	$\bar{X}$ = 83.7	$\bar{X}$ = 57.9
K	% Student Initiated	25.2	19.6	28.0
	% Student Response	74.8	80.4	72.0
		$\bar{X}$ = 92.9	$\bar{X}$ = 94.0	$\bar{X}$ = 78.4
1	% Student Initiated	21.8	22.6	19.6
	% Student Response	78.2	77.4	80.4
		$\bar{X}$ = 100.1	$\bar{X}$ = 92.0	$\bar{X}$ = 90.1
2	% Student Initiated	19.4	21.9	19.3
	% Student Response	80.6	78.1	80.7
		$\bar{X}$ = 79.4	$\bar{X}$ = 72.4	$\bar{X}$ = 78.2
3	% Student Initiated	19.5	24.4	18.1
	% Student Response	80.5	75.6	81.9
		$\bar{X}$ = 73.8	$\bar{X}$ = 84.6	$\bar{X}$ = 75.4
4	% Student Initiated	24.3	25.9	20.2
	% Student Response	75.7	74.1	79.8
				$\bar{X}$ = 86.7
5	% Student Initiated	*	*	18.2
	% Student Response	*	*	81.8
				$\bar{X}$ = 60.7
6	% Student Initiated	*	*	21.8
	% Student Response	*	*	78.2

Table 20

Mean Proportion of Teacher Behaviors by Category, Program and Grade (Pooled Across Languages)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 205.7	$\bar{X}$ = 324.6	$\bar{X}$ = 186.8
K	% Explain	12.4	18.8	22.5
	% Question	23.8	23.4	20.8
	% Command	18.9	21.3	27.9
	% Modeling	9.2	6.8	4.6
	% Feedback	20.8	19.0	19.6
	% Other	3.9	3.0	0.2
	% Monitoring	11.0	7.7	4.5
		$\bar{X}$ = 256.2	$\bar{X}$ = 293.7	$\bar{X}$ = 247.9
1	% Explain	16.9	21.9	19.5
	% Question	25.1	23.5	25.9
	% Command	22.4	20.8	23.2
	% Modeling	4.3	5.6	3.7
	% Feedback	23.2	19.5	20.9
	% Other	2.1	1.7	1.0
	% Monitoring	6.1	7.0	5.8
		$\bar{X}$ = 298.4	$\bar{X}$ = 284.9	$\bar{X}$ = 240.0
2	% Explain	19.3	20.4	14.4
	% Question	24.8	23.8	23.0
	% Command	21.2	20.0	25.9
	% Modeling	3.8	4.1	2.7
	% Feedback	21.1	20.8	23.3
	% Other	0.7	1.2	0.6
	% Monitoring	9.1	9.7	10.1
		$\bar{X}$ = 242.0	$\bar{X}$ = 269.0	$\bar{X}$ = 241.0
3	% Explain	21.2	23.3	19.3
	% Question	22.0	21.7	24.1
	% Command	18.9	20.3	22.0
	% Modeling	4.4	3.8	3.0
	% Feedback	18.3	16.8	19.8
	% Other	0.6	0.5	0.5
	% Monitoring	14.5	13.6	11.4

Table 20 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 244.8	$\bar{X}$ = 302.5	$\bar{X}$ = 233.3
4	% Explain	21.0	26.2	18.8
	% Question	21.2	20.5	28.4
	% Command	19.7	18.8	19.5
	% Modeling	4.7	3.4	1.3
	% Feedback	15.7	17.4	21.0
	% Other	0.1	0.0	0.7
	% Monitoring	17.6	13.7	10.5
				$\bar{X}$ = 230.4
5	% Explain	*	*	18.8
	% Question	*	*	22.5
	% Command	*	*	22.8
	% Modeling	*	*	2.5
	% Feedback	*	*	20.9
	% Other	*	*	0.5
	% Monitoring	*	*	12.0
				$\bar{X}$ = 192.7
6	% Explain	*	*	18.6
	% Question	*	*	19.2
	% Command	*	*	21.7
	% Modeling	*	*	3.1
	% Feedback	*	*	15.9
	% Other	*	*	0.5
	% Monitoring	*	*	20.9

Table 21

Mean Proportion of Teacher Behaviors in English by Category, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 175.7	$\bar{X}$ = 196.3	$\bar{X}$ = 20.6
K	% Explain	14.3	19.3	14.6
	% Question	27.2	23.2	19.8
	% Command	20.3	24.9	31.7
	% Modeling	10.6	7.8	8.4
	% Feedback	23.5	21.4	25.4
	% Other	4.0	3.3	0.0
	% Monitoring	0.0	0.0	0.0
		$\bar{X}$ = 231.0	$\bar{X}$ = 184.5	$\bar{X}$ = 81.6
1	% Explain	17.8	23.2	20.3
	% Question	27.0	25.0	26.8
	% Command	23.4	23.3	25.1
	% Modeling	4.7	6.0	8.4
	% Feedback	25.1	20.9	19.0
	% Other	2.0	1.7	0.4
	% Monitoring	0.0	0.0	0.0
		$\bar{X}$ = 266.7	$\bar{X}$ = 197.9	$\bar{X}$ = 76.6
2	% Explain	21.4	21.8	12.9
	% Question	27.7	25.7	23.4
	% Command	23.0	22.6	27.0
	% Modeling	4.2	4.8	5.3
	% Feedback	23.0	23.4	31.5
	% Other	0.7	1.7	0.1
	% Monitoring	0.0	0.0	0.0
		$\bar{X}$ = 204.5	$\bar{X}$ = 189.2	$\bar{X}$ = 108.2
3	% Explain	24.8	24.6	19.3
	% Question	26.0	26.2	27.1
	% Command	21.9	22.6	24.9
	% Modeling	5.3	6.0	5.5
	% Feedback	21.3	19.9	22.5
	% Other	0.8	0.8	0.8
	% Monitoring	0.0	0.0	0.0

Table 21 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 202.0	$\bar{X}$ = 250.3	$\bar{X}$ = 109.7
4	% Explain	25.1	30.8	20.7
	% Question	25.6	24.2	29.3
	% Command	24.6	20.9	24.2
	% Modeling	5.7	4.0	2.2
	% Feedback	19.0	20.1	23.3
	% Other	0.2	0.0	0.3
	% Monitoring	0.0	0.0	0.0
				$\bar{X}$ = 126.4
5	% Explain	*	*	20.5
	% Question	*	*	25.0
	% Command	*	*	26.1
	% Modeling	*	*	3.8
	% Feedback	*	*	24.1
	% Other	*	*	0.5
	% Monitoring	*	*	0.0
				$\bar{X}$ = 117.8
6	% Explain	*	*	23.3
	% Question	*	*	23.7
	% Command	*	*	27.3
	% Modeling	*	*	4.2
	% Feedback	*	*	20.7
	% Other	*	*	0.7
	% Monitoring	*	*	0.0

When Spanish is considered, slight grade level differences exist within program (see Table 22). The first difference is the steady decline in the frequency with which Spanish is used by teachers within early-exit and late-exit programs as grade level increases. As immersion strategy teachers rarely use Spanish, care should be taken when reviewing data for immersion strategy classrooms. The exceptionally low frequency ($\bar{X}$ = 0.0 to $\bar{X}$ = 7.1) of utterances in Spanish among immersion strategy teachers limits the statements one can make regarding teacher behavior when Spanish is used.

ERIC

Table 22

Mean Proportion of Teacher Behaviors in Spanish by Category, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 4.3	$\bar{X}$ = 98.2	$\bar{X}$ = 155.3
K	% Explain	4.3	16.3	23.6
	% Question	25.7	38.5	22.0
	% Command	44.0	17.5	27.9
	% Modeling	1.4	5.8	4.1
	% Feedback	19.3	19.5	22.4
	% Other	5.4	2.3	0.1
	% Monitoring	0.0	0.0	0.0
		$\bar{X}$ = 5.5	$\bar{X}$ = 89.4	$\bar{X}$ = 151.1
1	% Explain	12.7	22.5	20.0
	% Question	25.0	26.8	27.4
	% Command	26.5	19.9	25.7
	% Modeling	2.8	7.3	3.0
	% Feedback	24.3	20.9	23.3
	% Other	8.6	2.7	0.4
	% Monitoring	0.0	0.0	0.0
		$\bar{X}$ = 7.1	$\bar{X}$ = 70.1	$\bar{X}$ = 142.3
2	% Explain	22.6	19.4	16.2
	% Question	19.2	26.1	23.7
	% Command	31.6	20.6	30.3
	% Modeling	3.1	10.9	3.0
	% Feedback	23.5	22.0	25.2
	% Other	0.0	1.0	0.6
	% Monitoring	0.0	0.0	0.0
		$\bar{X}$ = 4.4	$\bar{X}$ = 51.6	$\bar{X}$ = 100.6
3	% Explain	16.5	19.7	22.3
	% Question	40.2	33.5	28.1
	% Command	31.5	21.3	23.4
	% Modeling	0.0	2.6	2.4
	% Feedback	11.8	22.0	23.6
	% Other	0.0	0.8	0.2
	% Monitoring	0.0	0.0	0.0

Table 22 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 0.0	$\bar{X}$ = 14.4	$\bar{X}$ = 84.2
4	% Explain	0.0	35.7	20.0
	% Question	0.0	8.8	34.6
	% Command	0.0	30.9	20.8
	% Modeling	0.0	1.9	0.3
	% Feedback	0.0	22.6	23.8
	% Other	0.0	0.0	0.3
	% Monitoring	0.0	0.0	0.0
				$\bar{X}$ = 78.4
5	% Explain	*	*	19.0
	% Question	*	*	31.1
	% Command	*	*	22.1
	% Modeling	*	*	0.7
	% Feedback	*	*	26.9
	% Other	*	*	0.3
	% Monitoring	*	*	0.0
				$\bar{X}$ = 30.7
6	% Explain	*	*	23.4
	% Question	*	*	32.3
	% Command	*	*	27.5
	% Modeling	*	*	0.9
	% Feedback	*	*	15.4
	% Other	*	*	0.6
	% Monitoring	*	*	0.0

Explanations in Spanish tend to increase with grade (4.3% to 16.5%) within the immersion strategy program. With the exception of third grade, Spanish appears to be used most by immersion strategy teachers to command students, especially in kindergarten (44.0%). Similarly, early-exit teachers command (17.5% to 30.9%) and explain (16.3% to 35.7%) more as grade level increases. After an initial increase in teacher modeling behaviors from kindergarten (5.8%) to second grade (10.9%) in early-exit classrooms, there is a sharp drop in modeling in grades three (2.6%) and four (1.9%). While this decline might reflect increased student ability in their oral Spanish skills, it more likely reflects the rapid decline in program and teacher focus on the development of Spanish language skills in the early-exit program after grade one. With one exception, late-exit teachers exhibit similar behaviors across grades. The exception is that when using Spanish, late-exit teachers in grades four to six tend to ask more questions (31.1% to 34.6%) than teachers in kindergarten through grade three (22.0% to 28.1%). This is encouraging in that it suggests that late-exit teachers are using the students' primary language for instruction as exhibited by the higher incidence of questioning behavior.

Overall, comparing programs within grade level reveals minor differences in the way Spanish is used. Except for third grade immersion strategy teachers, teachers across programs provide students with approximately the same amount of feedback. Early-exit kindergarten teachers ask more questions (38.5%) than do immersion strategy or late-exit teachers (25.7% and 22.0%, respectively). Noticeably, when using Spanish, both immersion strategy and late-exit kindergarten teachers command more (44.0% and 27.9%, respectively) than their early-exit counterparts (17.5%). Early-exit and late-exit first grade teachers provide almost twice as much explanation (22.5% and 20.0%, respectively) as do immersion strategy first grade teachers (12.7%).

Minor differences notwithstanding, teachers in the three programs tend to use English and Spanish in the same way. That is, teachers tend to make the same types of statements in both languages.

<u>Do teachers speak differently to limited-English-proficient students than to fluent-English-proficient and/or English-only speaking students?</u>

A major assumption underlying the instruction of second language learners is that teachers need to differentiate their speech to students who are learning English (i.e., LEP students) from those students who are native speakers of English (i.e., EO students) or second language learners who have acquired sufficient proficiency in

English to function as effectively as native speakers of English (i.e., FEP students) (Krashen, 1981; Terrell, 1981). To determine if this differentiation occurs, teacher/student interactions were divided into three groups—teacher and LEP-only; teacher and FEP/EO-only; and teacher and LEP/FEP and/or EO mix. This was possible in that the language proficiency of the student or students with whom a teacher spoke was coded for each teacher utterance. The mean proportion of teacher utterances was calculated by type of utterance, program, and grade for each of these three student groups. By comparing teacher behavior by grade within program with LEP-only (see Table 23), FEP/EO-only (see Table 24), and LEP/FEP and/or EO-only (see Table 25) students, we can determine whether teachers say different things to each of these three student groups.

A few observations are noteworthy. First, teachers across programs tend to say the same things to students regardless of student language proficiency (see Tables 23 and 24). Exceptions tend to occur only in kindergarten. Immersion strategy kindergarten teachers explain more to FEP/EO-only students (16.0%) than to their LEP-only students (9.6%). Early-exit kindergarten teachers provide more explanations to LEP-only students (20.2%) than to FEP/EO-only students (12.7%). However, FEP/EO-only students in early-exit kindergarten classrooms receive about 50% more feedback than do LEP-only students (29.5% vs. 19.5%). Late-exit kindergarten FEP/EO-only students receive approximately 50% more questions (34.0%) and feedback (30.1%) than do LEP-only students (25.2% and 21.7%, respectively). LEP-only kindergartners receive almost 50% more commands (28.2%) than their FEP/EO-only classmates (19.1%).

Table 23

Mean Proportion of Teacher Behaviors (English and Spanish) to Limited English Proficient Students by Type, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 110.3	$\bar{X}$ = 157.0	$\bar{X}$ = 113.9
K	% Explain	9.6	20.2	19.6
	% Question	26.8	25.5	25.2
	% Command	23.3	20.5	28.2
	% Modeling	6.3	4.8	2.4
	% Feedback	24.9	19.5	21.7
	% Other	3.9	3.3	0.2
	% Monitoring	5.2	6.4	2.7
		$\bar{X}$ = 168.0	$\bar{X}$ = 166.9	$\bar{X}$ = 129.4
1	% Explain	14.0	18.9	17.0
	% Question	28.5	28.5	31.0
	% Command	21.6	20.1	21.6
	% Modeling	2.9	3.8	2.0
	% Feedback	26.6	23.8	24.7
	% Other	2.1	1.6	1.0
	% Monitoring	4.1	3.3	2.6
		$\bar{X}$ = 185.6	$\bar{X}$ = 156.6	$\bar{X}$ = 134.5
2	% Explain	16.3	16.7	13.0
	% Question	29.8	28.1	26.1
	% Command	20.5	21.3	24.2
	% Modeling	2.1	2.1	2.0
	% Feedback	26.0	25.4	29.2
	% Other	0.6	1.3	0.7
	% Monitoring	4.6	5.1	4.8
		$\bar{X}$ = 131.8	$\bar{X}$ = 131.8	$\bar{X}$ = 122.7
3	% Explain	16.6	18.3	15.1
	% Question	28.6	28.1	30.0
	% Command	20.9	22.8	22.1
	% Modeling	2.4	1.4	1.2
	% Feedback	25.6	22.6	26.8
	% Other	0.7	0.7	0.3
	% Monitoring	5.1	6.0	4.5

Table 23 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 75.4	$\bar{X}$ = 137.1	$\bar{X}$ = 75.3
4	% Explain	14.1	21.1	12.7
	% Question	34.0	25.6	37.1
	% Command	22.9	23.2	18.3
	% Modeling	0.3	0.9	0.9
	% Feedback	25.9	23.7	28.5
	% Other	0.0	0.1	0.8
	% Monitoring	2.9	5.4	1.6
				$\bar{X}$ = 74.0
5	% Explain	*	*	10.7
	% Question	*	*	29.3
	% Command	*	*	23.0
	% Modeling	*	*	1.4
	% Feedback	*	*	30.8
	% Other	*	*	0.5
	% Monitoring	*	*	4.4
				$\bar{X}$ = 53.4
6	% Explain	*	*	13.1
	% Question	*	*	29.4
	% Command	*	*	24.3
	% Modeling	*	*	1.4
	% Feedback	*	*	26.2
	% Other	*	*	0.2
	% Monitoring	*	*	5.4

Table 24

Mean Proportion of Teacher Behavior to Fluent English Proficient and/or English-Only Speaking Students by Type, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 52.9	$\bar{X}$ = 76.1	$\bar{X}$ = 34.7
K	% Explain	16.0	12.7	16.1
	% Question	29.7	27.0	34.0
	% Command	23.3	23.7	19.1
	% Modeling	2.1	3.1	0.0
	% Feedback	25.3	29.5	30.1
	% Other	2.8	2.8	0.0
	% Monitoring	1.0	1.3	0.7
		$\bar{X}$ = 56.3	$\bar{X}$ = 48.6	$\bar{X}$ = 42.7
1	% Explain	11.8	17.1	18.3
	% Question	33.3	29.0	33.2
	% Command	20.2	21.4	20.5
	% Modeling	1.2	2.5	0.2
	% Feedback	31.3	27.2	26.0
	% Other	1.1	1.6	0.9
	% Monitoring	1.2	1.1	0.9
		$\bar{X}$ = 67.3	$\bar{X}$ = 56.8	$\bar{X}$ = 54.9
2	% Explain	14.1	13.2	9.6
	% Question	30.4	31.1	33.6
	% Command	17.7	21.8	20.8
	% Modeling	1.3	0.6	0.6
	% Feedback	30.6	30.9	33.1
	% Other	0.9	1.1	0.4
	% Monitoring	5.1	1.3	2.1
		$\bar{X}$ = 43.4	$\bar{X}$ = 71.6	$\bar{X}$ = 64.1
3	% Explain	14.9	17.2	11.5
	% Question	33.0	29.0	30.0
	% Command	19.6	20.8	25.1
	% Modeling	0.6	1.0	0.2
	% Feedback	27.3	26.1	28.6
	% Other	0.5	0.7	0.9
	% Monitoring	4.0	5.1	3.7

Table 24 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 60.4	$\bar{X}$ = 61.5	$\bar{X}$ = 76.4
4	% Explain	11.9	18.8	14.2
	% Question	25.1	26.1	35.9
	% Command	27.1	22.8	17.4
	% Modeling	0.0	0.1	0.3
	% Feedback	26.2	28.7	29.6
	% Other	0.5	0.0	0.8
	% Monitoring	9.2	3.4	2.0
				$\bar{X}$ = 91.2
5	% Explain	*	*	16.7
	% Question	*	*	29.3
	% Command	*	*	21.5
	% Modeling	*	*	0.6
	% Feedback	*	*	28.0
	% Other	*	*	0.2
	% Monitoring	*	*	3.7
				$\bar{X}$ = 82.2
6	% Explain	*	*	14.3
	% Question	*	*	29.3
	% Command	*	*	23.6
	% Modeling	*	*	0.5
	% Feedback	*	*	24.8
	% Other	*	*	0.6
	% Monitoring	*	*	6.9

Table 25

Mean Proportion of Teacher Behaviors to Mixed Groups of LEP, FEP, and/or EO Students by Type, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 140.4	$\bar{X}$ = 95.9	$\bar{X}$ = 73.3
K	% Explain	22.9	19.0	32.0
	% Question	11.4	16.7	6.2
	% Command	12.5	19.0	29.0
	% Modeling	16.3	14.7	13.7
	% Feedback	10.6	12.7	11.0
	% Other	2.4	2.6	0.0
	% Monitoring	24.0	15.2	8.1
		$\bar{X}$ = 105.6	$\bar{X}$ = 95.1	$\bar{X}$ = 66.2
1	% Explain	31.0	28.9	28.6
	% Question	14.8	11.8	15.8
	% Command	21.2	21.6	26.6
	% Modeling	8.6	9.2	7.1
	% Feedback	12.6	7.8	9.1
	% Other	1.7	2.1	0.3
	% Monitoring	10.1	18.6	12.4
		$\bar{X}$ = 99.9	$\bar{X}$ = 99.6	$\bar{X}$ = 68.6
2	% Explain	26.1	28.5	22.7
	% Question	16.0	14.6	11.4
	% Command	20.8	21.1	31.3
	% Modeling	7.5	7.8	3.6
	% Feedback	7.9	8.6	8.4
	% Other	0.9	0.3	0.4
	% Monitoring	20.8	19.1	22.2
		$\bar{X}$ = 106.5	$\bar{X}$ = 90.8	$\bar{X}$ = 84.1
3	% Explain	28.1	30.5	25.8
	% Question	11.9	10.1	10.7
	% Command	15.0	20.7	24.7
	% Modeling	8.0	7.5	6.3
	% Feedback	7.0	5.7	5.6
	% Other	0.9	0.0	0.2
	% Monitoring	29.0	25.5	26.7

ERIC

Table 25 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 109.0	$\bar{X}$ = 107.3	$\bar{X}$ = 72.6
4	% Explain	30.4	36.4	27.3
	% Question	9.1	9.2	13.5
	% Command	14.0	14.0	24.3
	% Modeling	10.6	7.5	2.6
	% Feedback	3.7	3.5	5.3
	% Other	0.0	0.0	0.4
	% Monitoring	32.3	29.4	26.7
				$\bar{X}$ = 82.5
5	% Explain	*	*	28.4
	% Question	*	*	10.7
	% Command	*	*	23.7
	% Modeling	*	*	6.7
	% Feedback	*	*	5.5
	% Other	*	*	0.8
	% Monitoring	*	*	24.2
				$\bar{X}$ = 84.5
6	% Explain	*	*	22.9
	% Question	*	*	7.2
	% Command	*	*	21.8
	% Modeling	*	*	5.9
	% Feedback	*	*	1.5
	% Other	*	*	0.0
	% Monitoring	*	*	40.8

The second most interesting finding is that while teachers across programs may not differentiate their speech between LEP-only and FEP/EO-only student groups, they pointedly speak differently when interacting with LEP, FEP and/or EO-only students mixed together (see Tables 23 through 25). Teachers across programs and grades consistently explain and model more often to mixed groups of students (Table 25) than to either LEP-only (Table 23) or FEP/EO-only (Table 24) student groups. LEP-only and FEP/EO-only student groups are questioned approximately twice as often and receive more feedback than mixed groups of students. Not surprisingly, teachers are more likely to engage in monitoring behavior with mixed groups of students than with LEP-only or FEP/EO-only groups.

In sum, the finding that teachers in all three programs tend to provide more explanations when speaking to mixed groups of students than when speaking to LEP-only or FEP/EO-only students may reflect the teacher's concern that he/she must take extra pains to ensure that the LEP student is following the discussion in a mixed setting. However, when LEP students are separated, presumably the teacher's statements to them can be directed more to their level of understanding, thereby not requiring extra explanation. Similarly, the observation that teachers in all three programs seem to prefer to ask more questions of students when they are separated by language status (LEP-only or FEP/EO-only) than when they are mixed also suggests sensitivity to the differential needs of students with varying language proficiency. Given the differences in language proficiency among students, certain types of questions (e.g., referential when speaking in Spanish and display when speaking in English) might be more appropriate for LEP students, and other types of questions for FEP/EO-only students (e.g., referential questions when using English). That teachers prefer to exhibit more modeling behavior with mixed groups of students is also reasonable in that in a modeling situation one can easily mix students, and, given the expected repetition, all students can participate equally. Similarly, teachers tend to monitor mixed groups of students in all three programs more than they do LEP-only or FEP/EO-only student groups. Monitoring suggests a somewhat passive instructional procedure. Although teachers do spend a noticeable portion of their time in passive instructional activity, when it does occur, it is more often with mixed student groups. This suggests that while teaching behaviors tend to be passive in mixed student groups, teacher behavior tends to be somewhat more active when students are separated by language proficiency. Thus, teachers in all three programs exhibit some sensitivity to the different language proficiency levels of their students.

What do teachers talk about?

Whenever teachers initiated a conversation with students, the focus of their conversation was coded as form (i.e., how language is used), concept (i.e., expressing an idea), or non-verbal (e.g., when gestures were used to communicate with students or when the teacher was monitoring the students). The data in Table 26 suggest that, across programs and grade levels, the majority of teacher-initiated conversations with students focus on concepts (55.9% to 82.4%). Teacher initiations that focus on form range from 7.5% to 23.5%. This pattern of verbal initiations also holds true whether teachers use English or Spanish (see Tables 27 and 28). Teachers also use gestures or other non-verbal behaviors to communicate with students (8.8% to 29.0%). In the absence of any specific data which would suggest an "ideal" mix of types of focus, this pattern appears to be somewhat desireable for language learning. Some argue that in the early stages of language development, language learning is facilitated when the focus is on the communication of meaning (i.e., what is said) rather than on how language should be used (i.e., how it is said) (Penfield & Roberts, 1958; Krashen, 1981).

When language is not considered (see Table 26), minor grade level differences are noted within program. Albeit somewhat uneven, immersion strategy and early-exit teachers initiate fewer interactions with students on form as grade level increases. Unfortunately, most of this decrease appears to reflect a concomitant increase in use of non-verbal behaviors to initiate student interactions. This is not ideal from a language learning perspective, as it appears to limit the opportunity for "comprehensible input" and "output." Again, while not an even trend, after limited interactions on form in kindergarten (8.8%), late-exit teachers begin to focus more on form in grades one (16.1%), two (12.0%), and three (12.8%), only to subside in grade four (7.5%), then slowly increase again in grades five (9.9%) and six (10.9%). What is most striking in these observations for each of the programs is the lack of a systematic approach to the focus on form and concept. The

ERIC

data seem to reflect a lack of understanding by teachers (and the curriculum they use) of when and how much to focus on form as opposed to concept.

Table 26

Mean Proportion of Teacher Initiation by Focus, Program, and Grade

Grade		Program Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 122.1	$\bar{X}$ = 148.2	$\bar{X}$ = 82.0
K	% Form	23.5	20.1	8.8
	% Concept	55.9	64.9	82.4
	% Non-Verbal	20.6	15.0	8.8
		$\bar{X}$ = 141.9	$\bar{X}$ = 151.1	$\bar{X}$ = 129.2
1	% Form	19.3	21.6	16.1
	% Concept	68.8	65.1	73.8
	% Non-Verbal	12.0	13.3	10.1
		$\bar{X}$ = 158.0	$\bar{X}$ = 159.6	$\bar{X}$ = 149.9
2	% Form	14.5	13.9	12.0
	% Concept	70.2	69.7	72.3
	% Non-Verbal	15.3	16.5	15.7
		$\bar{X}$ = 149.1	$\bar{X}$ = 149.4	$\bar{X}$ = 139.4
3	% Form	14.8	14.0	12.8
	% Concept	61.8	63.6	70.1
	% Non-Verbal	23.4	22.4	17.1
		$\bar{X}$ = 164.8	$\bar{X}$ = 165.0	$\bar{X}$ = 130.6
4	% Form	18.2	14.8	7.5
	% Concept	56.8	62.5	74.9
	% Non-Verbal	25.0	22.7	17.6
				$\bar{X}$ = 146.6
5	% Form	*	*	9.9
	% Concept	*	*	71.0
	% Non-Verbal	*	*	19.1
				$\bar{X}$ = 137.3
6	% Form	*	*	10.9
	% Concept	*	*	60.1
	% Non-Verbal	*	*	29.0

When language is considered, teachers use English and Spanish differently. Simply looking at the mean frequency, one finds that both immersion strategy and early-exit teachers typically use English when initiating a conversation with students across grade levels (see Tables 27 and 28). Among late-exit teachers, the frequency of teacher-initiated conversation in English increases as grade level increases, to the point that by sixth grade more than three times as many teacher initiations are in English ($\bar{X}$ = 76.4) than in Spanish ($\bar{X}$ = 18.4).

While there is an uneven tendency for the proportion of teacher initiations in English and Spanish dealing with concepts to increase as grade level increases in each program, when grade levels within programs are examined, minor language differences emerge.

Although immersion strategy teachers across all grades seldom use Spanish, when it is used, they focus almost exclusively on concepts (91.4% to 100%) as compared to when they use English (69.9% to 82.7%). Early-exit teachers appear to focus more on form in Spanish in kindergarten (25.6%), first grade (33.5%), and second grade (28.5%) than in English (21.5%, 21.8%, and 15.2%, respectively). This pattern suggests teacher understanding that in the early stages of acquiring a second language, language learning is facilitated when interactions emphasize concept over form. As second language learners already have a foundation in their primary language, it would be appropriate and necessary for teachers to focus more often on form in the primary language in the hopes of facilitating its development. This pattern is reversed in grades three and four, wherein early-exit teachers focus more on form when using English (20.2% and 19.4%, respectively) than when using Spanish (15.2% and 12.1%, respectively). This change presumably reflects changes in early-exit curriculum and increased concern over how English is used in preparation for mainstreaming and a decreased emphasis over the development of primary language skills.

ERIC

Table 27

Mean Proportion of Teacher Initiations in English by Focus, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 94.8	$\bar{X}$ = 83.9	$\bar{X}$ = 16.2
K	% Form	30.1	21.5	15.8
	% Concept	69.9	78.5	84.2
		$\bar{X}$ = 121.7	$\bar{X}$ = 90.6	$\bar{X}$ = 42.4
1	% Form	22.5	21.8	19.9
	% Concept	77.5	78.2	80.1
		$\bar{X}$ = 132.5	$\bar{X}$ = 102.3	$\bar{X}$ = 46.6
2	% Form	17.3	15.2	16.7
	% Concept	82.7	84.8	83.3
		$\bar{X}$ = 113.5	$\bar{X}$ = 94.3	$\bar{X}$ = 58.5
3	% Form	19.5	20.2	21.9
	% Concept	80.5	79.8	78.1
		$\bar{X}$ = 123.7	$\bar{X}$ = 124.2	$\bar{X}$ = 61.1
4	% Form	22.5	19.4	11.9
	% Concept	77.5	80.6	88.1
				$\bar{X}$ = 73.8
5	% Form	*	*	16.7
	% Concept	*	*	83.3
				$\bar{X}$ = 76.4
6	% Form	*	*	17.1
	% Concept	*	*	82.9

Table 28

Mean Proportion of Teacher Initiations in Spanish by Focus, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 2.4	$\bar{X}$ = 43.9	$\bar{X}$ = 62.3
K	% Form	5.8	25.6	8.3
	% Concept	94.2	74.4	91.7
		$\bar{X}$ = 4.5	$\bar{X}$ = 45.4	$\bar{X}$ = 73.3
1	% Form	8.6	33.5	16.0
	% Concept	91.4	66.5	84.0
		$\bar{X}$ = 3.3	$\bar{X}$ = 38.8	$\bar{X}$ = 86.2
2	% Form	8.5	28.5	13.6
	% Concept	91.5	71.5	86.4
		$\bar{X}$ = 2.8	$\bar{X}$ = 35.1	$\bar{X}$ = 52.2
3	% Form	0.0	15.2	14.9
	% Concept	100.0	84.8	85.1
		$\bar{X}$ = 0.0	$\bar{X}$ = 4.0	$\bar{X}$ = 43.3
4	% Form	0.0	12.1	5.5
	% Concept	0.0	87.9	94.5
				$\bar{X}$ = 48.4
5	% Form	*	*	3.8
	% Concept	*	*	96.2
				$\bar{X}$ = 18.4
6	% Form	*	*	3.9
	% Concept	*	*	96.1

Contrary to expectations, late-exit teachers tend to focus a relatively consistent proportion of their English interactions on form across grades (15.8% to 21.9%). As noted above, one would have expected a lower proportion of interactions focusing on form in kindergarten with a gradual increase in form as grade level increased. When using Spanish, the focus of late-exit teachers' interactions begins with a relatively low proportion of form in kindergarten (8.3%),

ERIC

slightly higher in first (16.0%), second (13.6%), and third (14.9%) grades, then sharply decreasing in grades four (5.5%), five (3.8%), and six (3.9%). While the higher levels of interactions that focus on form suggest attention to the development of primary language skills in the lower grades, the marked decrease in form in the upper elementary grades is of concern. Typically it is in these later grades wherein language skills are consolidated and expanded to more sophisticated levels. This pattern suggests that late-exit teachers may not be sufficiently developing their students' primary language skills.

What types of questions do teachers ask?

Questioning, the second most frequently occurring teacher behavior, was coded in three ways to describe the nature of information sought from the students. One count coded the frequency with which the student was asked for previously learned information (i.e., display questions--What is two times two? What happened to Humpty Dumpty?), for new information (i.e., referential questions--How many ways can you use a newspaper? How are you feeling today?), or for information to confirm/clarify what was said or understood (i.e., What did you say? Can you tell us in you own words?).

Overall, disregarding the language used, the patterns are consistent across programs and grade levels, with minor exceptions. In general, two to three times as many display questions are asked for every referential question (see Table 29). This pattern suggests that most teacher questioning consists of simple information recall. Research indicates that greater student achievement is realized in classrooms where higher-order, more cognitively demanding questions are more evident than are repetitive recall-type questions (Redfield & Rousseau, 1981). These patterns suggest that all three programs provide less demanding instructional environments than should be provided.

ERIC

Table 29

Mean Proportion of Teacher Questions by Type, Program, and Grade

Grade		Program	
	Immersion Strategy	Early-Exit	Late-Exit
	$\bar{X}$ = 46.0	$\bar{X}$ = 58.4	$\bar{X}$ = 33.6
K % Display	61.9	67.0	64.1
% Referential	32.6	31.5	33.2
% Clarification	5.4	1.5	2.6
	$\bar{X}$ = 64.7	$\bar{X}$ = 63.4	$\bar{X}$ = 59.8
1 % Display	65.4	63.5	71.8
% Referential	30.7	33.2	24.3
% Clarification	3.9	3.4	3.9
	$\bar{X}$ = 69.4	$\bar{X}$ = 65.5	$\bar{X}$ = 55.4
2 % Display	65.2	69.9	67.5
% Referential	32.1	27.8	28.9
% Clarification	2.8	2.3	3.6
	$\bar{X}$ = 54.2	$\bar{X}$ = 53.0	$\bar{X}$ = 55.5
3 % Display	60.3	57.6	61.0
% Referential	36.4	39.2	34.8
% Clarification	3.3	3.2	4.2
	$\bar{X}$ = 53.6	$\bar{X}$ = 55.6	$\bar{X}$ = 59.3
4 % Display	46.0	56.4	73.9
% Referential	52.8	40.8	22.3
% Clarification	1.3	2.8	3.7
			$\bar{X}$ = 52.2
5 % Display	*	*	66.9
% Referential	*	*	28.6
% Clarification	*	*	4.4
			$\bar{X}$ = 37.4
6 % Display	*	*	52.3
% Referential	*	*	43.3
% Clarification	*	*	4.3

Regardless of the language used, teachers exhibit the same behaviors in questioning students. Whether English or Spanish is used, teachers almost always ask display questions at least twice, if not three times, as often as referential questions (see Tables 30 and 31). There are four exceptions. When using English, fourth grade immersion strategy teachers ask slightly more referential questions (52.8%) as they do display questions (46.0%) (see Table 30). Similarly, among early-exit fourth grade and late-exit sixth grade teachers, slightly less than half of their questions are referential (40.8% and 41.6%, respectively) and slightly more than half are display questions (56.5% and 53.0%, respectively). When Spanish is used, third grade early-exit teachers use referential questions over half of the time (52.4%) (see Table 31). Late-exit sixth grade teachers use referential questions (47.9%) about as equally as they do display questions (49.0%). The frequency of questions asked in Spanish by immersion strategy teachers is too low to interpret its distribution ($\overline{X} = 1.4$ to $\overline{X} = 3.0$).

Table 30

Mean Proportion of Teacher Questions in English by Type, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 45.3	$\bar{X}$ = 37.0	$\bar{X}$ = 5.5
K	% Display	62.4	61.9	31.7
	% Referential	32.5	36.8	58.3
	% Clarification	5.2	1.3	10.0
		$\bar{X}$ = 62.3	$\bar{X}$ = 43.3	$\bar{X}$ = 24.5
1	% Display	65.8	59.3	62.2
	% Referential	30.4	37.2	32.2
	% Clarification	3.9	3.5	5.7
		$\bar{X}$ = 68.3	$\bar{X}$ = 50.2	$\bar{X}$ = 25.5
2	% Display	65.5	68.3	62.2
	% Referential	31.8	28.7	35.5
	% Clarification	2.7	2.9	2.4
		$\bar{X}$ = 53.5	$\bar{X}$ = 43.5	$\bar{X}$ = 27.7
3	% Display	60.8	58.9	62.7
	% Referential	31.8	28.7	35.5
	% Clarification	2.7	2.9	2.4
		$\bar{X}$ = 53.6	$\bar{X}$ = 54.5	$\bar{X}$ = 31.4
4	% Display	46.0	56.5	73.0
	% Referential	52.8	40.8	23.7
	% Clarification	1.3	2.8	3.3
				$\bar{X}$ = 30.3
5	% Display	*	*	66.8
	% Referential	*	*	27.6
	% Clarification	*	*	5.6
				$\bar{X}$ = 27.8
6	% Display	*	*	53.0
	% Referential	*	*	41.6
	% Clarification	*	*	5.3

ERIC

Table 31

Mean Proportion of Teacher Questions in Spanish by Type, Program, Grade

Grade		Program: Immersion Strategy	Program: Early-Exit	Program: Late-Exit
K		$\bar{X}$ = 1.8	$\bar{X}$ = 23.8	$\bar{X}$ = 29.9
	% Display	25.3	73.4	67.3
	% Referential	37.6	25.5	29.9
	% Clarification	37.1	1.1	2.8
1		$\bar{X}$ = 3.0	$\bar{X}$ = 24.3	$\bar{X}$ = 36.8
	% Display	33.9	71.5	72.1
	% Referential	56.5	25.4	24.6
	% Clarification	9.7	3.2	3.3
2		$\bar{X}$ = 2.4	$\bar{X}$ = 19.9	$\bar{X}$ = 35.7
	% Display	38.0	76.5	66.9
	% Referential	57.3	21.9	28.0
	% Clarification	4.7	1.6	5.0
3		$\bar{X}$ = 1.4	$\bar{X}$ = 15.0	$\bar{X}$ = 27.9
	% Display	0.0	44.6	63.9
	% Referential	100.0	52.4	32.6
	% Clarification	0.0	3.0	3.5
4		$\bar{X}$ = 0.0	$\bar{X}$ = 1.9	$\bar{X}$ = 27.1
	% Display	0.0	75.4	74.0
	% Referential	0.0	24.6	20.0
	% Clarification	0.0	0.0	6.0
5				$\bar{X}$ = 23.9
	% Display	*	*	61.9
	% Referential	*	*	34.9
	% Clarification	*	*	3.2
6				$\bar{X}$ = 10.3
	% Display	*	*	49.0
	% Referential	*	*	47.9
	% Clarification	*	*	3.2

Overall, the data suggest that, regardless of the language used, typically teachers ask primarily information recall questions, rather

ERIC

than more analytical, cognitively demanding questions (i.e., When was the Guttenberg press invented? vs. Why was the invention of the Guttenberg press so important?).

<u>What do students say?</u>

Whenever students initiated a conversation, their behaviors were coded into one of four categories: (a) asking a question (What time do we have lunch?); (b) making a free comment (the teacher mentions that Christopher Columbus was from Italy; student pipes up and says, "My Uncle Paolo is from Italy."); (c) non-verbal (student makes no verbal statement, but makes a gesture to initiate a conversation, such as raising his hand for the teacher's attention); and (d) other (verbal statements used to initiate a conversation that could not be coded in one of the other three categories, such as "Mrs. Juarez, kick the ball with your right foot.").

Without considering language, across programs and grades, students initiate a conversation by asking a question or by making a free comment (see Table 32). Students seldom use gestures in any program to initiate a conversation with their teachers. In each program, the frequency of free comments decreases while that of asking questions increases as grade level increases. There appear to be only two consistent program differences. The first is that, across grades, immersion strategy and early-exit students consistently initiate more interactions with their teachers than do late-exit students. Second, immersion strategy and early-exit students tend to initiate about twice as many conversations with teachers using non-verbal cues than do late-exit students. This is interesting, in that one might have predicted that with the greater use of the students' primary language in the classroom that late-exit students would have been more forthcoming with their teachers. The lower frequency of late-exit student-initiated interactions may also reflect more structured late-exit classrooms than immersion strategy or early-exit classrooms. Further, while immersion strategy students are permitted to use their primary language in the classroom and while early-exit teachers also use their students' primary language in the classroom, it would appear that the language use pattern of the teacher is a more powerful predictor of student language use. That is, "Do as I say, not as I do," may not be any more effective in influencing child (i.e., student) verbal behavior than it is in influencing other social behavior. Nonetheless, the low frequency of student- initiated interactions, especially among late-exit classrooms, would suggest less than optimal language learning environments in all three programs. As noted before, language learning is facilitated when there is ample opportunity to produce language in meaningful interactions, that the student has the opportunity to initiate as well as to respond to another speaker. Once again, the pattern that emerges is one in which students are passive language learners.

Table 32

Mean Proportion of Student Initiating Behaviors by Type, Program, and Grade

Grade		Program: Immersion Strategy	Program: Early-Exit	Program: Late-Exit
		$\bar{X}$ = 20.7	$\bar{X}$ = 23.0	$\bar{X}$ = 19.7
K	% Ask Question	28.1	35.2	39.7
	% Free Comment	58.9	53.6	60.3
	% Non-Verbal	11.5	11.2	0.0
	% Other	1.5	0.0	0.0
		$\bar{X}$ = 22.0	$\bar{X}$ = 26.4	$\bar{X}$ = 17.6
1	% Ask Question	33.6	43.1	52.1
	% Free Comment	54.7	43.4	41.5
	% Non-Verbal	10.3	11.8	5.5
	% Other	1.4	1.7	1.0
		$\bar{X}$ = 22.0	$\bar{X}$ = 22.1	$\bar{X}$ = 16.3
2	% Ask Question	39.6	45.4	50.3
	% Free Comment	47.9	41.5	42.3
	% Non-Verbal	11.6	9.9	6.5
	% Other	0.9	3.3	1.0

Table 32 (continued)

Grade		Program Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 18.3	$\bar{X}$ = 23.1	$\bar{X}$ = 15.9
3	% Ask Question	49.9	44.9	48.3
	% Free Comment	35.4	41.6	42.3
	% Non-Verbal	14.8	12.9	7.1
	% Other	0.0	0.6	2.4
		$\bar{X}$ = 21.9	$\bar{X}$ = 27.8	$\bar{X}$ = 16.6
4	% Ask Question	56.2	44.9	65.2
	% Free Comment	25.9	40.8	31.2
	% Non-Verbal	16.7	14.3	3.0
	% Other	1.1	0.0	0.6
				$\bar{X}$ = 14.7
5	% Ask Question	*	*	48.8
	% Free Comment	*	*	45.9
	% Non-Verbal	*	*	4.0
	% Other	*	*	1.3
				$\bar{X}$ = 15.0
6	% Ask Question	*	*	58.4
	% Free Comment	*	*	29.5
	% Non-Verbal	*	*	10.4
	% Other	*	*	1.7

When language is considered, students initiating conversations in all three programs tend to ask more questions and to make fewer free comments in both English and Spanish as grade level increase (see Tables 33 and 34). Although late-exit students use Spanish less and English more as grade level increases, proportionately more of their Spanish interactions are initiated by questions than their English interactions. No other clear patterns emerge when student initiating behaviors are examined by language.

ERIC

Table 33

Mean Proportion of Student Initiating Behaviors in English by Type, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 15.9	$\bar{X}$ = 14.8	$\bar{X}$ = 6.0
K	% Ask Question	31.3	42.2	41.4
	% Free Comment	66.6	57.8	58.6
	% Non-Verbal	0.0	0.0	0.0
	% Other	2.1	0.0	0.0
		$\bar{X}$ = 18.6	$\bar{X}$ = 16.1	$\bar{X}$ = 6.2
1	% Ask Question	38.4	46.9	51.5
	% Free Comment	60.1	52.0	47.8
	% Non-Verbal	0.0	0.0	0.0
	% Other	1.5	1.1	0.7
		$\bar{X}$ = 18.2	$\bar{X}$ = 15.1	$\bar{X}$ = 6.6
2	% Ask Question	45.3	51.3	52.7
	% Free Comment	53.8	45.5	45.7
	% Non-Verbal	0.0	0.0	0.0
	% Other	0.9	3.2	1.6
		$\bar{X}$ = 15.2	$\bar{X}$ = 16.1	$\bar{X}$ = 8.1
3	% Ask Question	59.7	49.5	51.8
	% Free Comment	40.3	49.2	44.9
	% Non-Verbal	0.0	0.0	0.0
	% Other	0.0	1.3	3.3
		$\bar{X}$ = 16.2	$\bar{X}$ = 21.1	$\bar{X}$ = 10.4
4	% Ask Question	65.0	52.4	67.5
	% Free Comment	33.8	47.6	31.6
	% Non-Verbal	0.0	0.0	0.0
	% Other	1.2	0.0	0.8
				$\bar{X}$ = 10.7
5	% Ask Question	*	*	49.8
	% Free Comment	*	*	48.5
	% Non-Verbal	*	*	0.0
	% Other	*	*	1.7

Table 33 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
				$\bar{X}$ = 11.1
6	% Ask Question	*	*	66.2
	% Free Comment	*	*	30.9
	% Non-Verbal	*	*	0.0
	% Other	*	*	2.9

<u>What do students say when they respond to teacher initiations</u>?

Student responses to teacher initiations were coded into one of nine categories. Four of them are the same as those defined earlier for student initiations: asking questions, making a free comment, non-verbal and other. The remaining categories are: (a) repetition (student response to a teacher statement modeling how something is said, usually occurring during a drill); (b) expected response (student attempts to give the single "right" answer in response to a question, such as "What is two times two?" or "Who was the first president of the United States?," although the response may actually be incorrect); (c) free response (student responds to an open-ended question, such as "What would you do with a million dollars?"); (d) no response (the student, not hearing the teacher, does not acknowledge the teacher's initiation bid); and (e) listening (the student is listening to a teacher's presentation).

When language is not considered, consistently across programs and grades, the most prevalent student responses are expected responses and non-verbal (see Table 35). The high frequency of expected responses (22.5% to 38.0%) reflects that display questions are one of the most frequent ways in which teachers initiate conversations with students. The high frequency of non-verbal student responses (37.0% to 47.3%) along with listening (3.9% to 14.5%) and no response (2.4% to 7.0%) underscores the passiveness of students' interactions with teachers

ERIC

(and by implication their learning environment for language and cognitive development). From a language development perspective, we noted earlier that information recall does not facilitate the acquisition of higher order language skills as much as more active referential questioning. We also noted earlier that it is important that students produce language if they are to develop language. The two patterns of student responses described above suggest that students in all three programs appear to have the opportunity for developing receptive language skills, but that the opportunities for developing productive language skills are restricted.

Table 34

Mean Proportion of Student Initiating Behaviors in Spanish by Type, Program, and Grade

Grade		Program Immersion Strategy	 Early-Exit	 Late-Exit
		$\bar{X}$ = 3.6	$\bar{X}$ = 7.2	$\bar{X}$ = 17.2
K	% Ask Question	61.8	48.6	39.4
	% Free Comment	38.2	51.4	60.6
	% Non-Verbal	0.0	0.0	0.0
	% Other	0.0	0.0	0.0
		$\bar{X}$ = 2.5	$\bar{X}$ = 7.9	$\bar{X}$ = 11.7
1	% Ask Question	19.8	46.9	53.8
	% Free Comment	80.2	51.1	45.3
	% Non-Verbal	0.0	0.0	0.0
	% Other	0.0	2.0	0.9
		$\bar{X}$ = 3.3	$\bar{X}$ = 7.6	$\bar{X}$ = 9.4
2	% Ask Question	57.9	37.3	58.4
	% Free Comment	42.1	59.1	41.1
	% Non-Verbal	0.0	0.0	0.0
	% Other	0.0	3.6	0.5

Table 34 (continued)

Grade		Program Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 1.2	$\bar{X}$ = 5.9	$\bar{X}$ = 7.1
3	% Ask Question	36.6	52.7	53.7
	% Free Comment	63.4	46.5	40.4
	% Non-Verbal	0.0	0.0	0.0
	% Other	0.0	0.8	5.9
		$\bar{X}$ = 1.0	$\bar{X}$ = 6.7	$\bar{X}$ = 6.5
4	% Ask Question	0.0	39.4	74.5
	% Free Comment	100.0	60.6	25.5
	% Non-Verbal	0.0	0.0	0.0
	% Other	0.0	0.0	0.0
				$\bar{X}$ = 3.4
5	% Ask Question	*	*	59.2
	% Free Comment	*	*	40.8
	% Non-Verbal	*	*	0.0
	% Other	*	*	0.0
				$\bar{X}$ = 1.6
6	% Ask Question	*	*	66.1
	% Free Comment	*	*	33.9
	% Non-Verbal	*	*	0.0
	% Other	*	*	0.0

Table 35

Mean Proportion of Student Responding Behaviors by Type, Program, and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 122.1	$\bar{X}$ = 192.1	$\bar{X}$ = 95.7
K	% Ask Question	1.1	0.3	0.5
	% Repetition	11.5	9.2	8.4
	% Expected Response	27.4	28.4	30.7
	% Free Response	6.2	6.7	10.1
	% Free Comment	3.7	2.3	1.9
	% Non-Verbal	40.3	41.8	42.1
	% No Response	5.4	4.3	2.4
	% Other	0.0	0.0	0.0
	% Listening	4.4	6.9	3.9
		$\bar{X}$ = 141.9	$\bar{X}$ = 166.9	$\bar{X}$ = 143.4
1	% Ask Question	1.1	1.0	1.1
	% Repetition	5.8	7.6	5.2
	% Expected Response	34.2	29.1	35.8
	% Free Response	7.7	8.3	6.0
	% Free Comment	2.6	2.7	2.0
	% Non-Verbal	37.3	38.7	37.9
	% No Response	4.8	4.6	4.8
	% Other	0.0	0.0	0.0
	% Listening	6.4	8.0	7.2
		$\bar{X}$ = 172.1	$\bar{X}$ = 166.5	$\bar{X}$ = 149.9
2	% Ask Question	1.1	0.9	0.6
	% Repetition	4.6	3.8	3.6
	% Expected Response	35.2	32.4	35.4
	% Free Response	7.4	5.8	6.3
	% Free Comment	2.0	1.7	2.5
	% Non-Verbal	39.2	41.3	41.0
	% No Response	3.8	5.9	4.4
	% Other	0.0	0.0	0.0
	% Listening	6.6	8.2	6.2

Table 35 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 149.1	$\bar{X}$ = 164.9	$\bar{X}$ = 146.7
3	% Ask Question	0.8	0.9	0.6
	% Repetition	3.8	2.3	2.3
	% Expected Response	30.2	25.0	33.3
	% Free Response	6.4	7.0	9.1
	% Free Comment	1.3	1.4	1.7
	% Non-Verbal	41.0	45.5	38.6
	% No Response	7.0	6.3	5.4
	% Other	0.0	0.0	0.0
	% Listening	9.6	11.7	8.9
		$\bar{X}$ = 164.8	$\bar{X}$ = 182.9	$\bar{X}$ = 137.0
4	% Ask Question	0.4	0.6	1.2
	% Repetition	2.2	1.1	1.7
	% Expected Response	22.5	26.9	38.0
	% Free Response	9.0	8.2	7.2
	% Free Comment	0.0	0.7	1.6
	% Non-Verbal	47.3	42.2	37.0
	% No Response	6.3	5.9	3.8
	% Other	0.0	0.0	0.0
	% Listening	12.2	14.5	9.6
				$\bar{X}$ = 146.6
5	% Ask Question	*	*	0.7
	% Repetition	*	*	2.2
	% Expected Response	*	*	37.1
	% Free Response	*	*	5.7
	% Free Comment	*	*	2.5
	% Non-Verbal	*	*	38.6
	% No Response	*	*	5.6
	% Other	*	*	0.0
	% Listening	*	*	7.7
				$\bar{X}$ = 137.3
6	% Ask Question	*	*	0.6
	% Repetition	*	*	2.2
	% Expected Response	*	*	24.7
	% Free Response	*	*	5.9
	% Free Comment	*	*	1.7
	% Non-Verbal	*	*	46.6
	% No Response	*	*	5.5
	% Other	*	*	0.0
	% Listening	*	*	12.8

When language is considered, the distribution of student responses changes dramatically. This largely reflects the coding procedures. All student behaviors are coded by language—English, Spanish, both English and Spanish, or none (no language). All non-verbal, no-response, and listening responses by definition are non-language. As such, when only those student responses in English-only or Spanish-only are coded, all non-language behaviors are omitted from the analysis, as are student responses wherein both English and Spanish are used. Consequently, Tables 36 and 37 do not include any non-language student responses or those wherein both English and Spanish are used concomitantly. While student response "other" is a verbal statement, its frequency is so low that it is dropped from the analysis. Consequently, only five of the nine student behaviors are included in Tables 36 and 37.

Table 36

Mean Proportion of Student Responding Behaviors in English by Type, Program, and Grade

Grade		Program	
	Immersion Strategy	Early-Exit	Late-Exit
	$\bar{X}$ = 52.7	$\bar{X}$ = 60.7	$\bar{X}$ = 11.7
K % Ask Question	2.1	0.4	1.8
% Repetition	23.8	18.8	36.3
% Expected Response	54.9	58.7	36.8
% Free Response	12.9	16.7	25.1
% Free Comment	6.3	5.3	0.0
	$\bar{X}$ = 71.4	$\bar{X}$ = 52.5	$\bar{X}$ = 26.7
1 % Ask Question	2.1	2.5	2.5
% Repetition	11.5	16.6	18.1
% Expected Response	67.3	54.2	66.3
% Free Response	14.5	19.6	9.6
% Free Comment	4.6	7.1	3.5

Table 36 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		X̄ = 88.0	X̄ = 56.7	X̄ = 27.6
2	% Ask Question	2.2	2.2	0.5
	% Repetition	9.7	8.9	7.4
	% Expected Response	69.5	69.6	67.9
	% Free Response	14.7	15.3	16.9
	% Free Comment	4.0	4.0	7.3
		X̄ = 62.9	X̄ = 47.4	X̄ = 34.5
3	% Ask Question	1.8	2.4	1.4
	% Repetition	9.8	8.8	5.8
	% Expected Response	70.4	64.9	69.5
	% Free Response	15.2	20.4	20.1
	% Free Comment	2.8	3.5	3.1
		X̄ = 55.7	X̄ = 68.3	X̄ = 38.0
4	% Ask Question	1.6	1.5	2.9
	% Repetition	6.6	2.7	5.6
	% Expected Response	66.1	71.6	74.8
	% Free Response	25.7	22.7	13.8
	% Free Comment	0.0	1.6	3.0
				X̄ = 44.6
5	% Ask Question	*	*	1.6
	% Repetition	*	*	5.7
	% Expected Response	*	*	74.9
	% Free Response	*	*	11.2
	% Free Comment	*	*	6.6
				X̄ = 38.3
6	% Ask Question	*	*	1.9
	% Repetition	*	*	7.4
	% Expected Response	*	*	66.6
	% Free Response	*	*	19.0
	% Free Comment	*	*	5.1

Table 37

Mean Proportion of Student Responding Behaviors in Spanish by Type, Program, and Grade

Grade		Immersion** Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 2.9	$\bar{X}$ = 28.7	$\bar{X}$ = 41.5
K	% Ask Question	13.8	1.7	1.0
	% Repetition	0.0	18.2	14.0
	% Expected Response	49.6	65.0	62.2
	% Free Response	18.5	12.4	18.1
	% Free Comment	18.1	2.6	4.7
		$\bar{X}$ = 2.1	$\bar{X}$ = 31.7	$\bar{X}$ = 46.6
1	% Ask Question	6.7	3.2	2.2
	% Repetition	9.0	15.2	8.0
	% Expected Response	41.8	63.8	74.6
	% Free Response	36.7	14.0	11.6
	% Free Comment	5.8	3.8	3.6
		$\bar{X}$ = 3.5	$\bar{X}$ = 23.6	$\bar{X}$ = 50.2
2	% Ask Question	8.1	1.7	1.6
	% Repetition	8.4	12.3	7.6
	% Expected Response	55.7	73.2	73.1
	% Free Response	25.9	9.3	12.3
	% Free Comment	1.8	3.5	5.5
		$\bar{X}$ = 1.2	$\bar{X}$ = 18.1	$\bar{X}$ = 32.0
3	% Ask Question	0.0	1.9	1.1
	% Repetition	0.0	3.1	3.4
	% Expected Response	11.7	77.3	74.2
	% Free Response	76.6	17.1	17.2
	% Free Comment	11.7	0.7	4.2
		$\bar{X}$ = 1.0	$\bar{X}$ = 3.9	$\bar{X}$ = 25.8
4	% Ask Question	0.0	5.2	2.2
	% Repetition	0.0	0.0	0.4
	% Expected Response	0.0	73.8	77.6
	% Free Response	100.0	5.2	13.3
	% Free Comment	0.0	15.7	6.5

Table 37 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
				$\bar{X}$ = 25.8
5	% Ask Question	*	*	3.3
	% Repetition	*	*	0.9
	% Expected Response	*	*	67.0
	% Free Response	*	*	18.0
	% Free Comment	*	*	10.8
				$\bar{X}$ = 9.4
6	% Ask Question	*	*	0.6
	% Repetition	*	*	1.0
	% Expected Response	*	*	58.2
	% Free Response	*	*	28.0
	% Free Comment	*	*	12.1

** The extremely low frequency renders the distribution for the program meaningless.

Table 36 presents the data for student responses in English. Once again, across programs and grades the most frequent responses are expected responses (36.8% to 74.9%). Also, expected responses tend to increase (albeit somewhat unevenly) as grade level increases. Additional grade level differences are noted within each program. Students repeat less and provide more free responses as grade level increases within both immersion strategy and early-exit programs. Late-exit students also repeat less as grade level increases. However, unlike students in immersion strategy and early-exit programs, the frequency of free responses among late-exit students fluctuates widely as grade level increases.

When student responses in Spanish are examined, as before, the most frequent student responses are expected responses. While the proportion of expected responses tends to increase with grade, the pattern is inconsistent, fluctuating from grade to grade within program. Comparing the frequency of student responses in English with those in Spanish by grade clearly illustrates that students in immersion strategy programs almost always use English, and documents how students in both early-exit and late-exit programs decrease their use of Spanish while increasing their use of English. In fact, the frequency of student responses in Spanish is so low that the distribution is not useful. Program comparisons by grade suggest that students in early-exit and late-exit programs generally exhibit similar behavior patterns when using Spanish.

In sum, the pattern of student responses suggests a less than optimum environment for developing oral language skills in that students in all three programs have limited opportunities to produce language and, when they do produce language, it is limited to information recall. This occurs whether English or Spanish is being used.

How often do students initiate a conversation with their teachers?

Consistently across programs and grades, the majority of student utterances (72.0% to 81.9%) are in response to teacher initiations (see Table 38). Once again, this underscores the limited opportunities students have to produce language that is not limited to information recall.

Table 38

Mean Proportion of Student Utterances by Type, Program, and Grade

Grade		Program Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 72.7	$\bar{X}$ = 83.7	$\bar{X}$ = 57.9
K	% Student Initiated	25.2	19.6	28.0
	% Student Response	74.8	80.4	72.0
		$\bar{X}$ = 92.9	$\bar{X}$ = 94.0	$\bar{X}$ = 78.4
1	% Student Initiated	21.8	22.6	19.6
	% Student Response	78.2	77.4	80.4

Table 38 (continued)

Grade		Program		
		Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 100.1	$\bar{X}$ = 92.0	$\bar{X}$ = 90.1
2	% Student Initiated	19.4	21.9	19.3
	% Student Response	80.6	78.1	80.7
		$\bar{X}$ = 79.4	$\bar{X}$ = 72.4	$\bar{X}$ = 78.2
3	% Student Initiated	19.5	24.4	18.1
	% Student Response	80.5	75.6	81.9
		$\bar{X}$ = 73.8	$\bar{X}$ = 84.6	$\bar{X}$ = 75.4
4	% Student Initiated	24.3	25.9	20.2
	% Student Response	75.7	74.1	79.8
				$\bar{X}$ = 86.7
5	% Student Initiated	*	*	18.2
	% Student Response	*	*	81.8
				$\bar{X}$ = 60.7
6	% Student Initiated	*	*	21.8
	% Student Response	*	*	78.2

<u>Do teachers shelter their language through the use of realia</u>?

Realia refers to pictures, pantomime, models, etc.; any strategy or device that a teacher might use along with what is being said to help the second language learner understand the meaning of the verbal message. For example, the teacher points to a color chart and says, "This is the color orange," while pointing out the corresponding color. Or the teacher says, "run," and he/she begins running.

With two exceptions, realia is not widely used in any of the three programs (see Table 39). Immersion strategy kindergarten teachers use realia the most, in almost half of their interactions. Early-exit kindergarten teachers use realia next most frequently, in about one-quarter of their interactions with students. It is interesting to note that immersion strategy kindergarten teachers, who use English to a greater extent, also use realia almost twice as often (46.8%) as do early-exit kindergarten teachers (27.5%). Both immersion strategy and early-exit teachers decrease their use of realia in first grade (17.6% and 15.0%, respectively), and all but discontinue its use in subsequent grades. In contrast, little realia is used among late-exit kindergarten teachers (0.5%), and is used somewhat by first grade teachers (14.4%). As late-exit kindergarten teachers do not tend to use much English, they do not use realia. These patterns appear to be reasonable for each program. As immersion strategy and early-exit kindergarten teachers use large amounts of English, they need to adjust or shelter their English in such a way as to make it more comprehensible to their students. As late-exit kindergarten teachers tend to use little English, they may not have the need to shelter their English. The subsequent reduction in realia by first grade immersion strategy and early-exit teachers may reflect an increase in student skills such that realia may not be needed as often as it was in kindergarten. By the same token, the increase in use of realia by first grade late-exit teachers may reflect their increase in the use of English to the point where they need to shelter it for their students. However, in the absence of data documenting the ideal level at which realia should be used, one can only wonder if the use patterns noted here are too abrupt to be truly effective over time. The overall low frequency of realia precludes us from determining the effects of varying levels of its use on achievement.

Table 39

Mean Proportion of Teacher Use of Realia by Program and Grade

Grade		Program: Immersion Strategy	Program: Early-Exit	Program: Late-Exit
		$\bar{X}$ = 142.7	$\bar{X}$ = 165.9	$\bar{X}$ = 98.9
K	% Used	46.8	27.5	0.5
	% Not Used	53.2	72.1	99.5
	% Unknown	0.0	0.4	0.0
		$\bar{X}$ = 163.9	$\bar{X}$ = 175.0	$\bar{X}$ = 145.0
1	% Used	17.6	15.0	14.4
	% Not Used	82.4	85.0	85.4
	% Unknown	0.0	0.0	0.2
		$\bar{X}$ = 178.2	$\bar{X}$ = 180.7	$\bar{X}$ = 166.2
2	% Used	7.1	7.6	[illegible]
	% Not Used	92.9	92.2	97.5
	% Unknown	0.0	0.2	0.0
		$\bar{X}$ = 167.5	$\bar{X}$ = 170.3	$\bar{X}$ = 154.5
3	% Used	2.0	2.4	0.9
	% Not Used	97.9	97.6	99.0
	% Unknown	0.0	0.0	0.1
		$\bar{X}$ = 186.7	$\bar{X}$ = 190.1	$\bar{X}$ = 146.4
4	% Used	0.1	0.6	8.3
	% Not Used	99.9	99.2	91.7
	% Unknown	0.0	0.1	0.0
				$\bar{X}$ = 161.3
5	% Used	*	*	0.5
	% Not Used	*	*	99.5
	% Unknown	*	*	0.1
				$\bar{X}$ = 152.4
6	% Used	*	*	0.2
	% Not Used	*	*	99.8
	% Unknown	*	*	0.0

Observational Context

There is substantial evidence documenting how the physical and social context of the environment affects human behavior (Ramirez, 1981). Consequently, it is important that key elements of the milieu in which the observations are made are examined for program and grade level differences. Should differences occur, they should be considered in the interpretation of the language observational data. For this discussion, two features of the classroom environment are described: classroom activities and content area.

Across what classroom activities were observations made?

Whenever an exchange was recorded between the teacher and student(s), the type of classroom activity in which the exchange was taking place was coded. Classroom activities were coded into one of eight categories: (a) presentation (students are listening to a presentation of material or directions from the teacher); (b) discussion (students are talking about appropriate content area); (c) seatwork (students are engaged in seatwork such as workbook, writing, silent reading, or other sanctioned non-academic activities such as puzzles, etc.); (d) drillwork (students are engaged in oral repetition); (e) listening (students are attending to audio equipment such as a tape recorder, radio, or phonograph); (f) interim (students are waiting for teacher's attention, sharpening a pencil, or lining up for recess); (g) reading (students are reading aloud); and (h) other (students are engaged in activities other than those described above, e.g., watching a filmstrip, working at the chalkboard, doing a science experiment, not fooling around).

Consistently across programs and grades, discussion is one of the major classroom activities, followed by seatwork (see Table 40). Nonetheless, some program differences by grade level are noted. In each program, the proportions of presentation, seatwork and reading activities increase, while drill, interim, and other activities decrease as grade level increases. Program differences tend to occur largely in kindergarten. Students are provided with more discussion and interim activities in early-exit (45.4% and 16.5%, respectively) and late-exit (41.8% and 28.0%, respectively) classrooms than in immersion strategy classrooms (31.0% and 13.7%, respectively). More seatwork and drills are provided in immersion strategy (21.2% and 12.1%, respectively) and early-exit (19.9% and 5.7%, respectively) than in late-exit (14.7% and 3.9%, respectively) classrooms. More discussion is also provided in immersion strategy and early-exit first grade classrooms than those in late-exit first grade classrooms. These few exceptions aside, activities tend to be fairly consistent across programs and grades.

Table 40

Mean Proportion of Classroom Activities by Program and Grade

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 142.7	$\bar{X}$ = 165.9	$\bar{X}$ = 98.9
K	% Presentation	5.5	4.1	1.3
	% Discussion	31.0	45.4	41.8
	% Seatwork	21.2	19.9	14.7
	% Drill	12.1	5.7	3.9
	% Listening	4.3	2.7	2.7
	% Interim	13.7	16.5	28.0
	% Reading	1.0	0.2	2.5
	% Other	11.2	5.5	5.1
		$\bar{X}$ = 163.9	$\bar{X}$ = 175.0	$\bar{X}$ = 145.0
1	% Presentation	6.6	6.8	10.3
	% Discussion	43.6	41.8	35.5
	% Seatwork	22.2	24.5	24.5
	% Drill	4.0	4.8	4.7
	% Listening	2.0	2.1	0.7
	% Interim	11.1	9.5	14.8
	% Reading	6.0	5.9	4.3
	% Other	4.4	4.8	5.3
		$\bar{X}$ = 178.2	$\bar{X}$ = 180.7	$\bar{X}$ = 166.2
2	% Presentation	6.9	7.5	4.7
	% Discussion	40.0	41.4	41.1
	% Seatwork	27.7	24.6	32.4
	% Drill	3.6	2.4	2.3
	% Listening	2.4	2.9	0.6
	% Interim	10.2	9.3	10.7
	% Reading	7.5	8.7	5.1
	% Other	1.7	3.1	3.1
		$\bar{X}$ = 167.5	$\bar{X}$ = 170.3	$\bar{X}$ = 154.5
3	% Presentation	6.7	8.6	7.2
	% Discussion	32.1	33.6	40.5
	% Seatwork	32.5	31.9	28.7
	% Drill	2.2	1.4	2.0
	% Listening	2.8	3.2	1.9
	% Interim	10.9	12.1	10.3
	% Reading	9.6	6.6	7.5
	% Other	3.2	2.7	1.9

ERIC

Table 40 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 186.7	$\bar{X}$ = 190.1	$\bar{X}$ = 146.4
4	% Presentation	8.0	8.2	9.3
	% Discussion	34.5	37.3	40.0
	% Seatwork	35.9	36.1	27.1
	% Drill	0.7	0.6	1.0
	% Listening	3.4	2.8	0.3
	% Interim	8.4	6.8	10.8
	% Reading	3.4	6.2	6.1
	% Other	5.6	2.0	5.4
				$\bar{X}$ = 161.3
5	% Presentation	*	*	4.9
	% Discussion	*	*	44.1
	% Seatwork	*	*	26.3
	% Drill	*	*	2.2
	% Listening	*	*	1.8
	% Interim	*	*	9.4
	% Reading	*	*	9.8
	% Other	*	*	1.5
				$\bar{X}$ = 152.4
6	% Presentation	*	*	6.2
	% Discussion	*	*	29.7
	% Seatwork	*	*	41.3
	% Drill	*	*	2.9
	% Listening	*	*	2.7
	% Interim	*	*	11.3
	% Reading	*	*	4.7
	% Other	*	*	1.2

Across what content areas were observations made?

Whenever an exchange between a teacher and student was recorded, content area was coded into one of seven areas: (a) reading; (b) language arts; (c) math; (d) social studies; (e) non-academic (e.g., art, music); (f) other academic (e.g., science, computer science); and (g) procedures (e.g., lining up for recess, collecting lunch money, taking attendance). Observations were made across all these areas in each program and grade; program differences by grade are noted (see Table 41). Most of the differences between programs occur in kindergarten. More observations were made during reading in immersion strategy (12.9%) and late-exit (17.3%) classrooms than in early-exit classrooms (6.0%). In addition, more observations were made during math, non-academic, and other academic areas in late-exit classrooms than in immersion strategy or early-exit classrooms. In contrast, more observations were made during language arts in early-exit (46.5%) and immersion strategy (30.5%) classrooms than in late-exit classrooms (16.0%). These differences are sufficiently large that they are taken into account in the achievement analysis.

Table 41

Mean Proportion of Observations by Content Area, Program, and Grade

Grade		Program Immersion Strategy	Early-Exit	Late-Exit
		$\bar{X}$ = 142.7	$\bar{X}$ = 165.9	$\bar{X}$ = 98.9
K	% Reading	12.9	6.0	17.3
	% Language Arts	30.5	46.5	16.0
	% Math	12.9	18.3	23.4
	% Social Studies	4.9	1.8	1.4
	% Non-Academic	14.9	9.9	17.6
	% Other Academic	6.4	0.4	15.5
	% Procedures	17.6	17.1	8.8
		$\bar{X}$ = 163.9	$\bar{X}$ = 175.0	$\bar{X}$ = 145.0
1	% Reading	31.1	27.5	11.5
	% Language Arts	26.8	28.0	33.1
	% Math	13.2	17.5	19.0
	% Social Studies	5.1	4.8	4.0
	% Non-Academic	4.8	6.4	8.1
	% Other Academic	6.5	4.5	11.2
	% Procedures	12.7	11.4	12.9
		$\bar{X}$ = 178.2	$\bar{X}$ = 180.7	$\bar{X}$ = 166.2
2	% Reading	22.4	26.9	15.7
	% Language Arts	31.9	26.8	33.9
	% Math	18.0	17.1	19.8
	% Social Studies	3.8	5.2	4.8
	% Non-Academic	6.7	6.5	2.5
	% Other Academic	7.4	8.0	11.8
	% Procedures	9.9	9.6	11.5
		$\bar{X}$ = 167.5	$\bar{X}$ = 170.3	$\bar{X}$ = 154.5
3	% Reading	25.2	21.4	29.5
	% Language Arts	28.1	25.5	26.3
	% Math	18.9	16.6	19.2
	% Social Studies	5.8	7.9	5.1
	% Non-Academic	6.3	9.6	6.0
	% Other Academic	6.8	8.1	7.1
	% Procedures	9.0	10.8	6.7

Table 41 (continued)

Grade		Immersion Strategy	Program Early-Exit	Late-Exit
		$\bar{X}$ = 186.7	$\bar{X}$ = 1[illegible]0.1	$\bar{X}$ = 146.4
4	% Reading	22.2	22.0	19.4
	% Language Arts	21.2	27.1	35.6
	% Math	18.5	13.5	17.5
	% Social Studies	10.1	10.2	8.6
	% Non-Academic	9.0	8.8	5.1
	% Other Academic	12.7	12.9	5.0
	% Procedures	6.3	5.4	8.8
				$\bar{X}$ = 161.3
5	% Reading	*	*	18.3
	% Language Arts	*	*	26.0
	% Math	*	*	21.3
	% Social Studies	*	*	11.8
	% Non-Academic	*	*	1.1
	% Other Academic	*	*	13.8
	% Procedures	*	*	7.8
				$\bar{X}$ = 152.4
6	% Reading	*	*	17.4
	% Language Arts	*	*	28.6
	% Math	*	*	15.7
	% Social Studies	*	*	7.5
	% Non-Academic	*	*	7.5
	% Other Academic	*	*	12.4
	% Procedures	*	*	11.0

1 5

ERIC

Implications

The function of the observational data is to document the implementation of the instructional models. Implementation is assessed by comparing the results of the observations against the theoretical underpinnings of each model. The discussion which follows focuses on the theoretical advantages inherent to the immersion strategy and bilingual program models. Bilingual programs which in theory afford students and teachers greater opportunities for the use of L1, Spanish in this case, should provide language and instruction which is more complex when L1 is used for instruction than when L2 is used for instruction. Several of the categories used to classify classroom discourse and activities are designed to note indirectly these behaviors and activities. Only those instances in which differences appear between program models are discussed.

Evidence for Greater Complexity of Instruction

Is there greater complexity in the language, content or context of utterances in bilingual late-exit programs as evidenced by:

1. students asking questions
2. students responding with a free response
3. less use of repetition
4. less use of drill
5. greater opportunities for discussion

Do students in the three programs differ in the percentage of questions asked?

Yes. It has been traditionally argued (Cohen, 1975, for example) that one of the inherent advantages of a program model which allows students to use their L1 is that students will be more likely to participate in instruction at a more complex level, particularly in the initial stages. The data collected for grades kindergarten through fourth grade suggest that in terms of percentage of questions asked in either language, immersion strategy students in grades kindergarten through the second grade ask proportionately fewer questions (28.1%, 33.6%, and 39.6%, respectively) than do students in either early-exit (35.2%, 43.1%, and 45.4%, respectively) or late-exit (39.7%, 52.1%, and 50.3%, respectively) programs (see Table 42). Most studies of L2 teaching strategies have not examined the role of student questions. As such, their relationship with achievement has not been established. In this study, the role of student questions is found not to be as important a factor contributing to student achievement as other factors (e.g., proportions of English and Spanish used).

Table 42

Proportion of Student Behaviors Indicating Complex Language Behavior, by Program and Grade

Student Behavior	Grade	Immersion %	Early-Exit %	Late-Exit %
% Ask Questions	K	28.1	35.2	39.7
	1	33.6	43.1	52.1
	2	39.6	45.4	50.3
	3	49.9	44.9	48.3
	4	56.2	44.9	65.2
% Free Comment	K	58.9	53.6	60.3
	1	54.7	43.4	41.5
	2	47.9	41.5	42.3
	3	35.4	41.6	42.3
	4	25.9	40.8	31.2
% Initiate	K	25.2	19.6	28.0
	1	21.8	22.6	19.6
	2	19.4	21.9	19.3
	3	19.5	24.4	18.1
	4	24.3	25.9	20.2

In general, students in the three programs do not appear to differ in the proportion of free comments made or interactions initiated.

__Are there differences in the nature of the activities that students are engaged in across programs? Do these differences indicate greater complexity of instructional content?__

No. The programs do not differ in the degree of complexity of instruction as defined through the relative mix (i.e., proportion) of repetition, discussion and drill activities made available to students (see Table 43). The proportion of discussion and drill is roughly comparable among the three programs from kindergarten through grade four.

Table 43

Proportion of Classroom Behaviors Indicating Complex Instructional Content, by Program and Grade

Classroom Behavior	Grade	Immersion %	Early-Exit %	Late-Exit %
% Repetition	K	11.5	9.2	8.4
	1	5.8	7.6	5.2
	2	4.6	3.8	3.6
	3	3.8	2.3	2.3
	4	2.2	1.1	1.7
% Discussion	K	31.0	45.4	41.8
	1	43.6	41.8	35.5
	2	40.0	41.4	41.1
	3	32.1	33.6	40.5
	4	34.5	37.3	40.0
% Drill	K	12.1	5.7	3.9
	1	4.0	4.8	4.7
	2	3.6	2.4	2.3
	3	2.2	1.4	2.0
	4	0.7	0.6	1.0

ERIC

Conclusion

- **Are the patterns of language use observed in the study's programs consistent with that of their respective instructional models?** Yes, for immersion strategy and early-exit classrooms. Somewhat, for late-exit classrooms. The proportion of English and Spanish used and the rate they are used across grades differs markedly by program and grade. These differences are consistent with the instructional model for the immersion strategy and early-exit programs. English is used almost exclusively for instruction at all grade levels. Early-exit teachers use English over two-thirds of the time in kindergarten and first grade, subsequently increasing its use to approximately three-fourths in grade two, over three-fourths in grade three, and almost always using English in grade four. The language use pattern among late-exit kindergarten through fourth grade classrooms is also consistent with this program's instructional model. However, the proportion of English used in the classroom continues to increase in grades five (63.6%) and six (80.3%), exceeding the 60% limit as defined for this instructional model. (Closer inspection of the data reveals that this reflects an increase in only one of the three late-exit programs. That is, there is some concern that this high level of English use (i.e., >60%) in these later grades might preclude sufficient use of the student's primary language to ensure its meaningful development for the late-exit program affected.)

- **Do students' use of English and Spanish differ by program and grade?** Yes. Students' use of language mirrors that of their teachers, differing by program and grade. Immersion strategy students almost always use English. Early-exit students increase their use of English from almost two-thirds in kindergarten to almost 100% in grade four. Late-exit students' English use increases at a much slower rate, from 9.3% in kindergarten to 80.3% in grade six.

- <u>What do teachers say</u>? Consistently across programs and grades, the most common statements made by teachers are explanation, question, command, and feedback. Only minor differences are noted by language. That is, teachers tend to make the same kinds of statements whether they use English or Spanish across programs and grades.

- <u>Do teachers speak differently to limited-English-proficient students than to fluent-English-proficient and/or English-only speaking students</u>? No. Teachers across programs and grades tend to say the same things to students regardless of student language proficiency when students are separated by language proficiency. Teachers across programs and grade levels speak differently to students who are mixed together by language proficiency. Teachers consistently explain and model more often to mixed groups of students (LEP/FEP/EO) than to either LEP-only or FEP/EO-only student groups. LEP-only and FEP/EO-only student groups receive approximately twice as many questions and receive more feedback than mixed groups of students. Finally, teachers are more likely to engage in monitoring behavior with mixed groups of students than with LEP-only or FEP/EO-only groups. These differences suggest teacher sensitivity in all three programs to the differential needs of students with varying levels of language proficiency.

- <u>What do teachers talk about</u>? Consistently across programs and grade levels, the majority of teacher-initiated conversations with students focus on concepts (i.e., ideas) rather than on form (i.e., how something is said). This is desirable from a second language learning perspective.

- <u>What do students say when they initiate a conversation</u>? Students initiate a conversation by asking questions or making free comments. This is consistent across programs, grades, and language.

ERIC

- <u>What do students say when they respond to teacher initiations?</u> When language is not considered, consistently across programs and grades, teacher initiations do not require students to produce language (i.e., non-verbal, no response, or listening). When students do provide a verbal response, they tend to provide expected responses (i.e., information recall). The high frequency of expected responses reflects that display questions are one of the most frequent ways in which teachers initiate conversations with students. Taken together, the high frequency of no language and expected responses underscores the passiveness of students' interactions with teachers, their primary source of language development in the classroom.

- <u>How often do students initiate a conversation with their teacher?</u> Consistently across programs and grades, the majority (>80%) of student language is in response to teacher initiations. Once again, this highlights the passive nature of language learning opportunities in classrooms specifically designed to meet the language learning needs of language minority students.

- <u>Do teachers shelter their use of language through realia?</u> There is some evidence of teacher attempts to adjust their language behavior through realia. However, use of realia appears to be limited to kindergarten by immersion strategy and early-exit teachers.

- <u>Across what classroom activities are observations available? Do they differ by program and grade?</u> Consistently across programs and grades, observations of teacher/student interactions are most often available during discussion, followed by seatwork activities. Minor program and grade level differences in the relative frequency of these and other activities are noted.

ERIC

- <u>Across what content areas are observations available?</u> Observations of teacher/student interactions are available across all content areas in each program and grade. Nonetheless, programs differ with respect to the amount of teacher/student interaction data available by content area.

In sum, the three programs in this study are clearly differentiated by their use of English and Spanish for instruction. These differences are consistent with their respective instructional models. Students use English and Spanish the same way as do their teachers. Teachers across programs, grades, and languages are consistent in their statements to students. They initiate the majority of interactions with students. Their usual statements to students follow the predictable teaching process—explanation, question, command, and feedback. Teachers are supportive of second language learning in that they say different things to student groups where LEP students are mixed in with FEP and EO students, and their statements to students tend to focus on what is said rather than how it is said. There is limited attempt to shelter language through realia among immersion strategy and early-exit teachers. However, data suggest that the language learning environment in classrooms across programs is limited. Students are not asked more cognitively- and language-demanding questions. Language learning is passive in that teachers do most of the talking, with students providing rote or non-langua[illegible] responses.

While observations of teacher/student interactions are available across activities and content areas by program and grade levels. These program and grade level differences are noted.

<u>Classroom Engaged Academic Time</u>

In Years 2, 3, and 4 of data collection, engaged time observations were gathered using one instrument which incorporated the measurement of academic engagement of student groups and classroom context. The question addressed was, "What is the level of task engagement of all students (both project study students and non-project students) in study classrooms?" Consequently, this section of the report discusses task engagement in reference to all students within study classrooms. Data collected in Years 2, 3 and 4 enabled the following questions to be answered.

<u>To what extent are students engaged in their assigned tasks? Does task engagement vary by program or grade?</u>

Students across programs and grades are engaged in their assigned tasks the majority of the time they are observed (86.5% to 96.3%) (see Table 44). Task engagement does not appear to differ by program or grade. That is, regardless of the program or grade, students attend to their assigned tasks most of the time.

Table 44

Mean Proportion Student Task Engagement, by Program and Grade (Classroom Engaged Academic Time Measure)

		Program		
Grade		Immersion Strategy	Early-Exit	Late-Exit
K	% Engaged	96.3	91.1	93.3
1	% Engaged	91.4	89.6	92.3
2	% Engaged	91.4	91.1	86.5
3	% Engaged	91.3	91.0	91.0
4	% Engaged	89.6	94.8	90.7
5	% Engaged	*	*	88.2
6	% Engaged	*	*	90.5

Do student grouping patterns vary by program and grade?

Yes. Data collected show similar patterns of student grouping strategies and classroom group activities among all teachers. Research (Wilson, et al., 1983) suggests that differences in student grouping and classroom task arrangements may account for variation in student engagement.

Teachers across grades and programs tend to cluster students in large groups (36.0% to 92.7%) (see Table 45). Nonetheless, slight grade level differences are noted within each program, particularly among late-exit classrooms. In contrast to other grade levels within the immersion strategy program, second grade teachers tend to use slightly more medium- (26.9%) and fewer large-size (39.3%) student groups. Fourth grade teachers use large-size groups the most (60.4%). After an initial start in kindergarten using mostly large-size student groups (51.3%) and a modest amount of small-size (20.7%) groups, early-exit first grade teachers increase their use of small-size groups (30.7%) and decrease their use of large-size student groups (36.0%). The distribution found in first grade reflects the grouping strategies of early-exit teachers in second and third grades. Late-exit teachers display the greatest variation across grades in student grouping strategies. There is a trend, albeit uneven, for late-exit teachers to increase the use of small groups (3.9% to 20.4%) as grade level increases. There is tremendous fluctuation in the use of large- and medium-size groups by late-exit teachers across grades. Roughly, the tendency is a decrease in the use of large-size groups from kindergarten to grade six (92.7% to 58.0%) and an increase in the use of medium-size groups (2.5%, kindergarten to 16.9%, sixth grade).

Table 45

Mean Proportion of Group Size, by Program and Grade (Classroom Engaged Academic Time Measure)

Gr.			Program: Immersion Strategy	Early-Exit	Late-Exit
K	% Individual	(1 Student)	8.4	10.1	0.9
	% Small	(2-5 Students)	29.6	20.7	3.9
	% Medium	(6-10 Students)	20.6	18.0	2.5
	% Large	(> 10 Students)	41.4	51.3	92.7
1	% Individual	(1 Student)	12.9	10.1	5.0
	% Small	(2-5 Students)	24.5	30.7	11.3
	% Medium	(6-10 Students)	17.5	23.2	20.0
	% Large	(> 10 Students)	45.2	36.0	63.7
2	% Individual	(1 Student)	10.2	10.3	5.7
	% Small	(2-5 Students)	23.6	29.1	17.9
	% Medium	(6-10 Students)	26.9	19.6	26.9
	% Large	(> 10 Students)	39.3	41.0	49.6
3	% Individual	(1 Student)	8.0	7.6	3.9
	% Small	(2-5 Students)	30.1	25.2	14.5
	% Medium	(6-10 Students)	14.7	24.6	20.7
	% Large	(> 10 Students)	47.2	42.6	60.9
4	% Individual	(1 Student)	4.9	5.8	5.9
	% Small	(2-5 Students)	22.3	23.4	10.3
	% Medium	(6-10 Students)	12.4	23.6	11.0
	% Large	(> 10 Students)	60.4	47.2	72.8
5	% Individual	(1 Student)	*	*	3.7
	% Small	(2-5 Students)	*	*	14.8
	% Medium	(6-10 Students)	*	*	26.6
	% Large	(> 10 Students)	*	*	55.0
6	% Individual	(1 Student)	*	*	4.8
	% Small	(2-5 Students)	*	*	20.4
	% Medium	(6-10 Students)	*	*	16.9
	% Large	(> 10 Students)	*	*	58.0

Program comparisons by grade indicate very similar grouping patterns among immersion strategy and early-exit teachers. Late-exit teachers consistently differ from their immersion strategy and early-exit colleagues at each grade. Teachers in immersion strategy and early-exit classrooms typically use large- and small-size groups. In contrast, late-exit teachers more characteristically use large- and medium-size student groups.

Do classroom group activities vary by program and grade?

Yes, but only slightly. Student groups appear to engage primarily in discussion and seatwork activities (see Table 46).

Table 46

Mean Proportion of Group Activity
by Type, Program, and Grade
(Classroom Engaged Academic Time Measure)

		Program		
Grade	Activity	Immersion Strategy	Early-Exit	Late-Exit
K	% Drillwork	7.9	3.0	0.0
	% Oral Reading	0.8	0.1	4.1
	% Seatwork	30.4	33.9	22.4
	% Listening	5.5	3.3	1.4
	% Teacher Presentation	4.9	3.1	0.0
	% Discussion	26.7	32.4	41.4
	% Other	7.2	3.6	4.8
	% Interim	16.7	20.5	25.9
1	% Drillwork	3.0	2.0	2.8
	% Oral Reading	4.4	3.1	4.5
	% Seatwork	36.3	42.5	33.9
	% Listening	3.3	2.7	1.5
	% Teacher Presentation	4.5	4.9	7.2
	% Discussion	30.7	29.2	28.1
	% Other	5.0	3.[illegible]	7.5
	% Interim	12.9	1[illegible]	14.5
2	% Drillwork	2.1	2.0	1.1
	% Oral Reading	3.6	4.0	5.0
	% Seatwork	38.6	40.4	35.0
	% Listening	2.4	2.4	1.3
	% Teacher Presentation	4.5	5.5	2.0
	% Discussion	31.2	30.9	34.7
	% Other	2.2	3.2	3.1
	% Interim	15.5	11.6	17.8

ERIC

Table 46 (continued)

Grade	Activity	Program: Immersion Strategy	Early-Exit	Late-Exit
3	% Drillwork	2.4	1.2	1.1
	% Oral Reading	5.8	5.3	6.4
	% Seatwork	38.4	36.6	33.6
	% Listening	3.9	5.2	2.[illegible]
	% Teacher Presentation	6.9	5.1	2.9
	% Discussion	22.0	25.2	33.6
	% Other	3.8	3.1	2.5
	% Interim	16.8	18.2	17.8
4	% Drillwork	0.0	0.7	0.6
	% Oral Reading	12.0	4.3	7.0
	% Seatwork	37.5	40.5	31.7
	% Listening	2.7	1.9	0.8
	% Teacher Presentation	2.4	3.9	6.4
	% Discussion	35.5	31.4	33.4
	% Other	3.6	1.6	6.6
	% Interim	6.5	15.8	13.5
5	% Drillwork	*	*	1.8
	% Oral Reading	*	*	8.8
	% Seatwork	*	*	27.6
	% Listening	*	*	1.5
	% Teacher Presentation	*	*	3.9
	% Discussion	*	*	35.2
	% Other	*	*	2.2
	% Interim	*	*	19.2
6	% Drillwork	*	*	0.8
	% Oral Reading	*	*	2.9
	% Seatwork	*	*	37.0
	% Listening	*	*	3.0
	% Teacher Presentation	*	*	4.3
	% Discussion	*	*	26.2
	% Other	*	*	1.5
	% Interim	*	*	24.2

While there tends to be an overall trend towards an increase in the proportion of discussion activities among immersion strategy teachers as grade level increases, there is a marked decrease in discussion activities in grade three. However, the proportion of discussion activities increases again in grade four. Coupled with the increase in discussion activities, the steady decline in the use of

ERIC

drillwork activities over time suggests a move towards a potentially more active language learning environment. (However, given the pattern of teacher/student interactions described earlier, discussion activities are not being fully utilized as well as they might be for language learning.)

In contrast, early-exit classroom teachers also exhibit a decline in the use of drillwork with a slight increase in the proportion of seatwork as grade increases. The proportion of discussion tends to remain about the same across grades, approximately one-third. From a language development perspective, these minor changes within the early-exit program simply signal a shift from one passive learning strategy to another.

Late-exit classrooms manifest the greatest grade to grade variation among the three instructional models. There is a noticeable increase in the proportion of seatwork and a decrease in discussion activity from kindergarten (22.4% and 41.4%, respectively) to first grade (33.9% and 28.1%, respectively). Thereafter, seatwork and discussion activity each comprise about one-third of total classroom activities. While there is a noteworthy reduction in the proportion of interim activities from kindergarten to first grade late-exit classrooms (25.9% to 14.5%), interim activities return to about the level found in kindergarten by the sixth grade (24.2%).

Grade level comparisons by program reveal only minor differences. Proportionately more discussion and interim activities are observed among late-exit (41.4% and 25.9%, respectively) and early-exit (32.4% and 20.5%, respectively) kindergarten classrooms than among immersion strategy kindergarten classrooms (26.7% and 16.7%, respectively). Proportionately more discussion is conducted by third grade late-exit teachers (33.6%) than by third grade immersion strategy (22.0%) or early-exit (25.2%) teachers. Finally, fourth grade early-exit teachers provide more seatwork activities (40.5%) than do early-exit and late-exit fourth grade teachers (37.5% and 31.7%, respectively).

Defining an active language learning environment as one in which a student is required to produce language spontaneously through original discourse would categorize the activities of drillwork, oral reading, seatwork, listening, and teacher presentation as passive language learning activities. Summing the proportion of these activities within each grade level by program, approximately half of the classroom activities in each program provide students with a passive language learning environment (43.4% to 57.4%). The exception is late-exit kindergarten, where approximately one-fourth of the classroom activities provide a passive language learning environment for students.

In sum, all students tend to be engaged most frequently in seatwork and discussion activities, with low to moderate involvement in drillwork, listening, interim or other activities. The distribution of these activities provides students with a less than ideal language learning environment. This is consistent for each of the three instructional models.

Across what content areas are observations of student task engagement conducted? Does this vary by program and grade?

With two exceptions, across programs and grades, the largest proportion of student task engagement observations was conducted during the language arts content area (20.1% to 43.4%) (see Table 47). Third grade late-exit and fourth grade immersion strategy classrooms are the exception, where more observations were conducted during reading (24.7% and 32.7%, respectively) than language arts (24.5% and 19.6%, respectively).

Table 47

Mean Proportion of Group Content Area,
by Program and Grade
(Classroom Engaged Academic Time Measure)

		Program		
Grade	Content Area	Immersion Strategy	Early-Exit	Late-Exit
K	% Reading	11.3	4.9	16.0
	% Language Arts	27.2	43.4	21.9
	% Math	12.8	17.8	21.1
	% Social Studies	6.8	1.6	0.0
	% Other Academic	8.1	9.0	16.6
	% Process/Structure	17.5	10.8	10.7
	% Non-Academic	16.4	12.4	13.7
1	% Reading	23.1	19.2	10.5
	% Language Arts	32.3	31.8	33.9
	% Math	14.6	18.7	17.6
	% Social Studies	3.9	4.8	5.8
	% Other Academic	7.9	4.9	10.5
	% Process/Structure	11.9	11.2	13.3
	% Non-Academic	6.3	9.4	8.4
2	% Reading	14.1	17.2	13.8
	% Language Arts	34.4	30.7	32.2
	% Math	18.1	17.8	17.1
	% Social Studies	3.5	5.4	4.5
	% Other Academic	7.2	8.5	9.6
	% Process/Structure	14.2	12.1	19.2
	% Non-Academic	8.6	8.2	3.8
3	% Reading	21.9	18.8	24.7
	% Language Arts	26.4	24.4	24.5
	% Math	18.8	14.7	18.4
	% Social Studies	6.2	8.2	5.7
	% Other Academic	7.0	8.6	6.2
	% Process/Structure	14.1	16.1	13.7
	% Non-Academic	5.7	9.3	6.7
4	% Reading	32.7	17.1	17.8
	% Language Arts	19.6	20.1	33.0
	% Math	17.1	13.0	17.8
	% Social Studies	8.0	10.0	7.4
	% Other Academic	9.5	12.3	6.6
	% Process/Structure	2.4	14.3	12.2
	% Non-Academic	10.7	13.1	5.4

Table 47 (continued)

Grade	Content Area	Program Immersion Strategy	 Early-Exit	 Late-Exit
5	% Reading	*	*	17.0
	% Language Arts	*	*	22.0
	% Math	*	*	20.2
	% Social Studies	*	*	10.4
	% Other Academic	*	*	10.2
	% Process/Structure	*	*	18.6
	% Non-Academic	*	*	1.5
6	% Reading	*	*	14.4
	% Language Arts	*	*	21.8
	% Math	*	*	16.9
	% Social Studies	*	*	6.7
	% Other Academic	*	*	12.6
	% Process/Structure	*	*	21.4
	% Non-Academic	*	*	6.2

There are minor grade level differences within the immersion strategy and early-exit programs. There is more substantial grade level variation within the late-exit program. While the increase is somewhat uneven, proportionately more student engagement observations were conducted during reading and less during language arts among immersion strategy and early-exit classrooms as grade level increases. Also increasing unevenly, more student task engagement observations were collected during process/structure as grade level increases in the late-exit program. There is a great deal of fluctuation across grades in the proportion of observations collected during language arts, reading, and non-academic classes in the late-exit program.

Minor program differences by grade level are limited to those in kindergarten. More student task engagement observations were collected during language arts and fewer during reading in kindergarten early-exit classrooms (43.4% and 4.9%, respectively) than in immersion strategy (27.2% and 11.3%, respectively) and late-exit (21.9% and 16.0%, respectively) classrooms. Moreover, approximately twice as many

ERIC

task engagement observations were collected in other academic areas from late-exit kindergarten classrooms (16.6%) than in either immersion strategy (8.1%) or early-exit (9.0%) classrooms.

Overall, student task engagement data were collected across comparable activities by program and grade.

Does group oral language use vary by program and grade?

Yes. Oral language use is defined as the oral language used among students within a group whether or not a teacher is present. That is, at the moment of observation, the oral language(s) used by any and all students in a given group is coded. Oral language use varies by program and grade consistent with the immersion strategy, early-exit and late-exit program models.

Student groups in immersion strategy classrooms use more English (56.9% to 79.2%) than those in early-exit (36.7% to 53.0%) or late-exit (10.2% to 46.0%) program classrooms (see Table 48). In contrast, late-exit student groups use Spanish more often (65.1% to 9.6%) than either early-exit (18.5% to 0.8%) or immersion strategy classrooms (3.0% to 0.8%). The proportions in which English and Spanish are used in the two bilingual programs are very different from each other. Whereas early-exit student groups use English about one-third and Spanish less than twenty percent of the time in kindergarten, the use of English slowly increases as grade level increases. Spanish use slowly decreases, and all but disappears in grade four (0.8%). In contrast, late-exit kindergarten students use English ten percent and Spanish two-thirds of the time. English use increases slowly in late-exit classrooms, so that it is not until grades three and four where English is used about one-third of the time (32.6% and 39.3%, respectively). Late-exit student groups use English about half of the time (46.0%) by grade six. This level of English use still does not equal the highest level of recorded use in either immersion strategy (79.2%) or early-exit (53.0%) fourth grade classrooms.

Table 48

Mean Proportion of Oral Language Use, by Program and Grade (Classroom Engaged Academic Time Measure)

Grade	Language	Program: Immersion Strategy	Program: Early-Exit	Program: Late-Exit
K	% English	65.3	36.7	10.2
	% Spanish	3.0	16.7	65.1
	% Both	9.6	12.2	13.6
	% No Oral Language	22.1	34.5	11.2
1	% English	57.8	38.4	25.1
	% Spanish	0.8	18.5	44.5
	% Both	7.8	10.9	8.7
	% No Oral Language	33.6	32.2	21.7
2	% English	56.9	42.9	16.2
	% Spanish	1.1	12.4	44.9
	% Both	6.5	11.4	13.5
	% No Oral Language	35.6	33.4	25.4
3	% English	56.9	48.4	32.6
	% Spanish	1.3	9.5	30.0
	% Both	5.7	10.9	14.3
	% No Oral Language	36.1	31.3	23.1
4	% English	79.2	53.0	39.3
	% Spanish	0.9	0.8	28.0
	% Both	4.4	9.3	9.5
	% No Oral Language	15.6	36.9	23.2
5	% English	*	*	39.5
	% Spanish	*	*	21.3
	% Both	*	*	14.1
	% No Oral Language	*	*	25.1
6	% English	*	*	46.0
	% Spanish	*	*	9.6
	% Both	*	*	16.2
	% No Oral Language	*	*	28.2

The most striking proportion in Table 48 is the amount of "no oral language" in each program by grade (11.2% to 36.9%). "No oral language" is coded whenever students are silently working on their

ERIC

assigned tasks. This variable is an indirect measure of the opportunity students have to produce or receive oral language, i.e., an indirect measure of oral language input and output. Roughly one-third of the time in immersion strategy and early-exit classrooms, students are in groups where there is no oral language. In contrast, only about one-fourth of the time do students in late-exit classrooms find themselves in groups where there is no oral language. These findings suggest limited oral language development opportunities for students in each of the three programs, especially for those in immersion strategy and early-exit classrooms.

In sum, while the use of English and Spanish varies in a manner consistent with the instructional model of each study program, overall, students are provided with little opportunity to produce language.

<u>Conclusion</u>

- <u>Are students engaged in their assigned tasks</u>? Yes. Consistently across programs and grades, students exhibit a high level of task engagement ($\geq$ 86.5%).

- <u>Do student grouping patterns vary by program and grade</u>? Somewhat. Consistently across programs and grades, students tend to be clustered into large groups (<u>>10 students/group</u>). However, as their second choice in grouping students, immersion strategy and early-exit teachers use small-size groups (<u>2-5 students</u>), whereas late-exit teachers use medium-size groups (<u>6-10 students</u>).

- <u>Do classroom group activities vary by program and grade</u>? No. Overall, across programs and grades, all students tend to be engaged primarily in discussion and seatwork activities. In each program, approximately half of the activities provided to students represent passive language learning activities.

- Is the distribution of task engagement observations across content areas the same by program and grade? Yes. While minor program differences are noted in kindergarten, overall task engagement observations are available from the same distribution of content areas with the largest proportion of observations conducted during the content area of language arts.

- Does group oral language vary by program and grade? Yes. The observed use of English and Spanish varies in a manner consistent with the instructional model of each program. Student groups in immersion strategy classrooms use more English than those in early-exit classrooms, who, in turn, use more English than student groups in late-exit classrooms. In contrast, late-exit student groups were observed to use more Spanish than their early-exit and immersion strategy counterparts.

In sum, students across programs and grades exhibit high levels of task engagement. Teachers in all three programs tend to group their students into large groups. Students in each program model use English and Spanish in a manner consistent with that of their respective instructional model. Task engagement observations were conducted across a comparable range of activities and content areas for each instructional program. There is some indication that the opportunities to produce language in most content areas are restricted, largely a result of passive classroom instructional activities.

ERIC

Teacher/Classroom Characteristics

Data which described the characteristics of each study classroom, school and/or district were collected by Aguirre International/SRA Technologies staff. Classroom data were collected annually. School and district information were collected during the first year of the study, and were updated as necessary. This information allowed for the following questions to be answered.

Are classrooms comparable between programs with respect to class size and composition?

Five hundred fifty-seven program classroom teachers are in the study, 166 in the immersion strategy program, 198 in the early-exit program, and 193 in the late-exit program. In each year of the study (see Tables 49 to 52), as well as overall (see Table 53), more early-exit and late-exit classrooms are involved in the study than immersion strategy classrooms. As a result, data collected at the classroom level by program must be weighted for program comparisons to be made.

ERIC

Table 49

Number of Program Classroom Teachers, by Program and Grade
Year 1: 1984-85

	Program					
	Immersion		Early-Exit		Late-Exit	
Grade	N	Weighted N	N	Weighted N	N	Weighted N
K	16	17.3	26	22.6	15	13.7
1	16	18.3	17	12.3	0	0.0
2	0	0.0	0	0.0	0	0.0
3	0	0.0	0	0.0	21	19.1
4	0	0.0	0	0.0	0	0.0
5	*	*	*	*	0	0.0
6	*	*	*	*	0	0.0
Total	32	35.6	43	35.0	36	32.8

Table 50

Number of Program Classroom Teachers, by Program and Grade
Year 2: 1985-86

	Program					
	Immersion		Early-Exit		Late-Exit	
Grade	N	Weighted N	N	Weighted N	N	Weighted N
K	13	14.7	12	14.5	7	8.1
1	25	27.4	27	26.5	24	21.6
2	11	12.9	17	12.3	0	0.0
3	0	0.0	0	0.0	7	8.1
4	0	0.0	0	0.0	18	15.9
5	*	*	*	*	0	0.0
6	*	*	*	*	0	0.0
Total	49	55.0	56	53.3	56	53.8

Table 51

Number of Program Classroom Teachers, by Program and Grade
Year 3: 1986-87

	Program					
	Immersion		Early-Exit		Late-Exit	
Grade	N	Weighted N	N	Weighted N	N	Weighted N
K	0	0.0	0	0.0	0	0.0
1	17	19.0	15	18.2	9	10.5
2	20	22.3	24	23.4	19	17.3
3	7	8.3	14	10.5	0	0.0
4	0	0.0	0	0.0	7	8.1
5	*	*	*	*	17	15.8
6	*	*	*	*	0	0.0
Total	44	49.6	53	52.0	52	51.7

Table 52

Number of Program Classroom Teachers, by Program and Grade
Year 4: 1987-88

	Program					
	Immersion		Early-Exit		Late-Exit	
Grade	N	Weighted N	N	Weighted N	N	Weighted N
K	0	0.0	0	0.0	0	0.0
1	0	0.0	0	0.0	0	0.0
2	19	21.2	13	15.7	8	9.3
3	17	18.8	21	20.9	19	16.7
4	5	5.3	12	8.7	0	0.0
5	*	*	*	*	7	8.1
6	*	*	*	*	15	13.2
Total	41	45.4	46	45.3	49	47.3

ERIC

Table 53

Number of Program Classroom Teachers, by Program and Study Year

	Program					
	Immersion		Early-Exit		Late-Exit	
Year	N	Weighted N	N	Weighted N	N	Weighted N
1984-85	32	35.6	43	35.0	36	32.8
1985-86	49	55.0	56	53.3	56	53.8
1986-87	44	49.6	53	52.0	52	51.7
1987-88	41	45.4	46	45.3	49	47.3
Total	166	185.7	198	185.6	193	185.6

Collapsed across years, the average number of students per classroom across grades is comparable among the immersion strategy (X = 23.1 to X = 25.8), early-exit (X = 22.6 to X = 26.4), and late-exit (X = 21.8 to X = 27.1) programs (see Table 54). Across programs, the largest study classroom has 38 students, while the smallest only has 8 students. The average number of target students per study classroom is also comparable across programs: immersion strategy (X = 8.4 to X = 16.0), early-exit (X = 6.6 to X = 17.0), and late-exit (X = 5.1 to X = 18.2) (see Table 55). Overall, the classroom with the largest number of target students has 31, while the classroom with the smallest number of target students only has 1. Understandably, as a result of attrition, the number of target students per classroom decreases as grade level increases.

Classroom composition as defined by proportion of LEP, FEP, and EO students varies by grade level within program, as well as between programs (see Table 56). Across grades, while the proportion of LEP students in immersion strategy classrooms remains at approximately three-fourths from kindergarten through third grade (73.2% to 80.9%), it drops to less than two-thirds in grade four (60.9%). The proportion of EO students ranges from 12.4% to 19.6% across grades. The proportion of FEP students increases, although somewhat unevenly, from

5.8% in kindergarten to 24.1% in grade four. Within the early-exit program, there is a steady and consistent decrease in the proportion of LEP students (76.8% to 50.5%) and an increase in FEP students (7.4% to 36.1%) as grade level increases. Similar to the immersion strategy program, the proportion of EO students ranges from 11.2% to 15.8%. In the late-exit program, there is a steady decrease in the proportion of LEP students (77.8% to 28.1%) and increase in FEP students (9.8% to 54.7%) as grade level increases, with the exception of grade two. Across grade levels, late-exit classrooms consistently have proportionately fewer EO students than classrooms in the other programs. All things being equal, the increase in the proportion of FEP students and decrease in the proportion of LEP students indirectly suggests some success among immersion strategy, early-exit, and late-exit programs in assisting students to acquire English skills sufficient to warrant them being reclassified by language proficiency. As noted earlier, the availability of native English speakers can enhance the acquisition of English by second language learners (assuming frequent and meaningful interactions between LEP and EO students).

Table 54

Mean, Minimum, and Maximum Class Sizes, by Program and Grade

	Program								
	Immersion			Early-Exit			Late-Exit		
Grade	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
K	25.6	14	32	26.4	17	34	27.1	20	33
1	23.1	10	31	24.0	12	32	21.8	8	33
2	23.8	10	38	23.3	10	33	22.3	14	30
3	25.3	11	34	23.9	8	35	23.6	11	33
4	25.8	18	29	22.6	9	31	24.6	15	33
5	*	*	*	*	*	*	22.6	9	33
6	*	*	*	*	*	*	25.3	15	35
Weighted number of responses	172			184			186		

Table 55

Mean, Minimum+, and Maximum Number of Target Students per Program Classroom, by Program and Grade

	Program								
	Immersion			Early-Exit			Late-Exit		
Grade	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
K	16.0	5	27	17.0	5	25	18.2	4	31
1	11.4	2	23	11.7	2	21	9.8	0	21
2	8.4	1	20	7.5	1	18	8.8	3	17
3	8.6	1	20	6.6	1	17	9.0	2	20
4	8.7	1	11	6.6	1	16	8.8	2	17
5	*	*	*	*	*	*	6.3	0	15
6	*	*	*	*	*	*	5.1	1	11
Weighted Number of Responses			186			186			186

+ Classrooms with zero target students are program classrooms where classroom level data were collected; however, all target students were transferred to other classrooms.

ERIC

Table 56

Mean Percentage of LEP, FEP, and EO Students,
by Program and Grade

Grade		Program Immersion %	 Early-Exit %	 Late-Exit %
K	LEP	74.6	76.8	77.8
	FEP	5.8	7.4	9.8
	EO	19.6	15.8	12.3
1	LEP	80.9	75.7	70.8
	FEP	6.7	10.3	19.7
	EO	12.4	14.0	9.5
2	LEP	73.2	72.6	79.1
	FEP	14.2	15.1	13.2
	EO	12.5	12.4	7.7
3	LEP	74.2	66.9	64.4
	FEP	12.0	21.9	26.5
	EO	13.8	11.2	9.0
4	LEP	60.9	50.5	52.4
	FEP	24.1	36.1	40.8
	EO	15.1	13.4	6.8
5	LEP	*	*	46.5
	FEP	*	*	47.3
	EO	*	*	6.2
6	LEP	*	*	28.1
	FEP	*	*	54.7
	EO	*	*	17.3
Weighted Number of responses		172	184	186

How was teacher oral language proficiency assessed?

Teacher oral language proficiency was assessed to establish a rating of primary teacher language skill for each target child in each program year. English and Spanish oral language abilities were evaluated in spring, 1986, for teachers participating in the study in 1984-85 and/or 1985-86. Teachers participating in the study in 1986-87 were evaluated in fall or spring, 1986-87, and teachers participating

ERIC

in the study in 1987-88 were evaluated in fall or spring, 1987-88. A detailed report of the teacher oral language assessment is available (Crespo, 1986).

Oral interview testing procedures developed by the Defense Language Institute/Language School (DLI/LS; December, 1982) were adapted for this study in order to assess English and Spanish language skills.

The DLI/LS assessment procedure is based on an in-person interview that assesses a person's language speaking ability when talking to a trained tester for a period of 10 to 40 minutes. The interviews are audiotaped and the tapes are retained for later review by a second rater. The resulting speech sample is rated on a 12-point scale ranging from 0 (for no practical ability to function in the target language) to 5 (for ability equivalent to that of an educated native speaker). Midpoints are assigned for ability which substantially surpasses the requirements for a given level but fails to sustain performance at the next higher level. The DLI/LS procedure measures general oral language abilities including diagnostic factors such as pronunciation, grammar, vocabulary, and oral interaction.

Definitions of Language Proficiency Levels. The criterion language skills required for ratings at each language level are defined in Table 57. The same criteria are applied to English and Spanish.

Table 57

Definitions of Language Proficiency Levels

NO PROFICIENCY:

Level 0 Unable to function in the spoken language. Oral production is limited to occasional isolated words. Essentially no communicative ability.

ELEMENTARY PROFICIENCY:

Level 1 Able to satisfy routine travel needs and minimum courtesy requirements. Can ask and answer questions on very familiar topics; within the scope of very limited language experience can understand simple questions and statements, allowing for slowed speech, repetition or paraphrase; speaking vocabulary inadequate to express anything but the most elementary needs; errors in pronunciation and grammar are frequent, but can be understood by Native Speaker (NS) used to dealing with foreigners attempting to speak the language; while topics which are "very familiar" and elementary needs vary considerably from individual to individual, any person at Level 1 should be able to order a simple meal, ask for shelter or lodging, ask and give simple direction, make purchases and tell time.

LIMITED WORKING PROFICIENCY:

Level 2 Able to satisfy routine social demands and limited work requirements. Can handle with confidence but not with facility most social situations including introductions and casual conversations about current events, as well as work, family and autobiographical information; can handle limited work requirements, needing help in handling any complications or difficulties; can get the essence of most conversations on non-technical subjects (i.e., topics which require no specialized knowledge) and has a speaking vocabulary sufficient to respond simply with some circumlocutions; accent, though often quite faulty, is intelligible; can usually handle elementary constructions quite accurately but does not have thorough or confident control of the grammar.

Table 57 (continued)

PROFESSIONAL WORKING PROFICIENCY:

Level 3 Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and information conversations on practical, social and professional topics. Can discuss particular interests and special fields of competence with reasonable ease; comprehension is quite complete for a normal rate of speech; vocabulary is broad enough that he rarely has to grope for a word; accent may be obviously foreign; control of grammar good; errors virtually never interfere with understanding and rarely disturb the NS.

FULL PROFESSIONAL PROFICIENCY:

Level 4 Able to use the language fluently and accurately on all levels normally pertinent to professional needs. Can understand and participate in any conversation within the range of own personal and professional experience with a high degree of fluency and precision of vocabulary; would rarely be taken for a NS, but can respond appropriately even in unfamiliar situations; errors of pronunciation and grammar quite rare; can handle informal interpreting from and into the language.

NATIVE OR BILINGUAL PROFICIENCY:

Level 5 Speaking proficiency equivalent to that of an educated native speaker (ENS). Has complete fluency in the language such that speech on all levels is fully accepted by ENS in all of its features, including breadth of vocabulary and idiom, colloquialisms and pertinent cultural references.

The organizing criterion is that of the educated native speaker (ENS). Testers compare speech productions to what an ENS would say in the same circumstance. A range of styles may be tested from common, every day speech to technical language, thus assessing the interviewee's ability to tailor language to particular audiences. Testers also look for evidence of linguistic mixtures in which the speaker incorporates words or phrases into his or her speech from two or more dialects or languages. Unless specifically requested, dialects are not tested. The focus is on standard languages which are not defined narrowly (e.g., Castillian Spanish), but admit regional variants of the educated standard. Thus, the standard educated Spanish spoken in Mexico, Puerto Rico, Cuba, etc., is acceptable.

Rating Procedures. Raters compare the teacher's speech sample to the level definitions described above and assign a global or overall score of the level at which the interviewee demonstrates mastery. In assigning a global score, raters consider the teacher's speaking and understanding ability to function in the target language (TL). More specifically, a person's linguistic behaviors are viewed from three different vantage points: (1) the functions the speaking behaviors represent; (2) the context in which they occur, including the content subsequently expressed; and (3) the accuracy with which these functions are accomplished. The three skill areas are considered by the rater twice: first, when eliciting the speech sample; and second, when rating what the speaker has successfully accomplished. Table 58 illustrates the inter-relationship of these three skills by oral proficiency level.

Table 58

Trisection of Oral Proficiency Levels

Oral Proficiency Level	Function	Context*	Accuracy
	(Tasks accomplished, attitudes expressed, tone conveyed)	(Topics, subjects areas, activities and jobs addressed)	(Acceptability, quality, and accuracy of message conveyed)
5	Functions in a manner that is equivalent to that of an (ENS).**	All subjects.	Performance equivalent to an ENS.

Table 58 (continued)

4	Able to tailor language to fit audience, counsel, persuade, negotiate, represent a point of view, and interpret informally for dignitaries.	All topics normally pertinent to professional needs.	Nearly equivalent to an ENS. Speech is extensive, precise, appropriate to every occasion with only occasional errors.
3	Can converse in formal and informal situations, resolve problem situations, deal with unfamiliar topics, provide explanations, describe in detail, offer supported opinions and hypothesize.	Practical, social, professional and abstract topics, particular interests, and special fields of competence.	Errors virtually never interfere with understanding and rarely disturb the NS.*** Only sporadic errors in basic structures.
2	Able to fully participate in casual conversations, can express facts, give instructions, describe, report on, and provide narration about current, past, and future activities.	Concrete topics such as own background, family, interests, work, travel and current events.	Understandable to an NS <u>not</u> used to dealing with foreigners; sometimes miscommunicates.
1	Can create with the language: ask and answer questions, participate in short conversations.	Everyday survival topics and courtesy requirements.	Intelligible to an NS used to dealing with foreigners.
0	No functional ability.	None.	Unintelligible.

* May be job specific
** ENS = Educated Native Speaker
*** NS = Native Speaker

How were teacher oral language proficiency interviewers trained?

Interviewers were trained during a five-day workshop. Training followed procedures outlined in an adapted version of the DLI/IS manual (Crespo, 1986). These procedures were supplemented by the use of demonstration anchor tapes. Trainees conducted two practice interviews during the workshop. All interviewers attained adequate reliability ($r \geq .85$) prior to completion of the training workshop; all ratings were within one-half point of the standard.

Are teachers comparable between programs with respect to oral language proficiency?

While the language proficiencies of classroom teachers in English by grade and program are roughly comparable, there are noteworthy differences in Spanish language proficiency between teachers in the immersion strategy and early-exit programs and teachers in the late-exit program (see Table 59). At each grade by program, teachers have the oral English language skills to teach effectively in English. Using the operational definitions of the twelve-point scale, it was posited that a minimum of 3.5 would be needed to allow a teacher to manipulate the language with sufficient facility to teach effectively. Mean ratings of English oral language proficiency of teachers across programs are above this minimum. However, on the average, the oral Spanish language proficiency of immersion strategy (1.3 to 2.8) and early-exit (2.2 to 3.2) teachers appears to be limited to engaging in basic social interaction (i.e., social chit-chat), and not sufficiently strong for effective instruction in Spanish. While the Spanish oral production skills may be too low for effective teaching by immersion strategy teachers, they might be sufficient to allow teachers to have the minimal receptive skills needed to understand what their students are saying when using Spanish. In contrast, the exceptionally low oral Spanish language skills among early-exit teachers raises serious concerns about their ability to teach effectively in Spanish with regard to the quality of Spanish instruction within the early-exit program. In contrast, late-exit teachers, on the average, have oral Spanish language skills sufficient to teach effectively in Spanish (3.9 to 4.6).

Teacher Characteristics

Of the 557 teachers participating in the study, questionnaires are available for 543 of them (see Table 60). These data represent a return rate of 97.5%. Data collected include teachers' personal characteristics, qualifications, teaching assignments, teaching practices, instructional philosophies, and their perceptions of their programs. These topics are discussed below.

ERIC

Table 59

Mean Ratings of Teachers' English and Spanish Proficiency+, by Program and Grade

Grade		Program Immersion	Early-Exit	Late-Exit
K	English	4.8	4.8	4.3
	Spanish	2.3	2.9	4.0
1	English	4.7	4.8	4.4
	Spanish	2.8	3.2	4.3
2	English	4.6	4.6	3.7
	Spanish	2.8	2.6	4.6
3	English	4.8	4.6	4.1
	Spanish	2.2	2.7	3.9
4	English	4.5	4.9	4.4
	Spanish	1.3	2.2	4.2
5	English	*	*	4.2
	Spanish	*	*	4.2
6	English	*	*	4.5
	Spanish	*	*	4.2
Weighted number of responses		171	172	163

+ Teacher speech samples were rated on a 12-point scale ranging from 0 (for no practical ability to function in the taught language) to 5 (for ability equivalent to that of an educated native speaker).

ERIC

Table 60

Number of Study Teachers Completing Teacher Interview, by Program and Grade

Grade	Immersion N	Immersion Weighted N	Early-Exit N	Early-Exit Weighted N	Late-Exit N	Late-Exit Weighted N
K	29	31.7	36	34.6	21	21.0
1	58	65.0	58	55.5	30	29.6
2	48	54.6	54	51.2	27	27.1
3	22	24.6	35	31.1	45	42.2
4	5	5.2	12	8.5	25	24.5
5	*	*	*	*	23	23.2
6	*	*	*	*	15	13.2
Total	162	181	195	181	186	181

ERIC

Are teachers comparable between programs with respect to demographics, language abilities, education, and experience?

Demographics. Teachers in the late-exit program are more likely to identify themselves as from an ethnic background similar to that of their students than teachers in either immersion strategy or early-exit programs (see Table 61). As one might have predicted from observed teacher oral language proficiency in Spanish, consistently more teachers in the late-exit program are Hispanic or Spanish (56.3% to 93.0%), as compared to teachers in either immersion strategy (18.1% to 67.5%) or early-exit (42.0% to 64.8%) programs. Conversely, more teachers in the immersion strategy (24.8% to 81.9%) and early-exit (34.5% to 58.0%) programs are white not of Hispanic origin, as compared to teachers in the late-exit program (2.8% to 25.4%).

Education and credentials. Late-exit teachers are better educated and have more specialized training to work with language minority children than teachers in immersion strategy or early-exit programs. Across grades, proportionately more teachers in the late-exit program (51.5% to 89.2%) have completed graduate education programs (i.e., master's degree or higher) than in either the immersion strategy (28.5% to 47.3%) or early-exit (13.2% to 42.3%) programs (see Table 62). While slightly more second and third grade immersion strategy teachers have a master's degree or doctorate as compared to their early-exit counterparts, overall, the educational level of teachers in these two programs is somewhat comparable. Similarly, across grade levels through grade four, late-exit teachers are more likely to have specialized training to work with language minority children, i.e., bilingual teaching (78.9% to 97.0%) and/or ESL (15.4% to 53.8%) credential, as compared to either immersion strategy (18.1% to 68.3%, and 14.8% to 41.6%, respectively) or early-exit teachers (19.8% to 82.5%, and 20.6% to 41.9%, respectively) (see Table 63). There is a general trend in both the immersion strategy and early-exit programs for proportionately fewer teachers to have specialized training as grade level increases. That is, the number of teachers having specialized training in both of these programs tends to be concentrated in the lower grades. At one level this distribution of teachers makes some sense given the limited number of such specially trained teachers. However, it does raise the question of the ability of these programs to maintain continuity of program quality as grade level increases for ongoing language minority students, and also for the ability of these programs to address the needs of new students entering their programs in the upper grades.

The shortage of specially certificated teachers in the fourth grade becomes apparent in the reporting of special coursework or in-service training on methods for addressing the needs of language minority children (see Table 64). Roughly two-thirds or more of the teachers report having received special training in ESL, second language acquisition, first language development, multi-cultural education, bilingual education, and/or language assessment. Once again, the proportion of immersion strategy and early-exit teachers having such training drops at grade four. In contrast, late-exit fourth grade teachers report relatively high proportions of such training, as well as teachers in grades five and six.

In sum, late-exit teachers consistently report having more education, special credentials, and more specialized coursework than either immersion strategy or early-exit teachers.

<u>Experience</u>. While teachers, across grades within and across programs, have a good deal of general teaching experience (X = 7.0 to 15.2 years), there are noticeable program and grade level differences in the amount of bilingual teaching experience (see Table 65). Late-exit teachers in kindergarten through second grade have more bilingual teaching experience (X = 7.2 to 8.2 years) than either their immersion strategy (X = 5.0 to 6.4 years) or early-exit (X = 6.3 to 6.7 years) counterparts. However, third and fourth grade immersion strategy teachers have more bilingual teaching experience (X = 6.8 and 6.4 years, respectively) than either late-exit (X = 6.6 and 5.9 years, respectively) or early-exit (X = 5.1 and 3.3 years, respectively)

teachers. In grades five and six, late-exit teachers reverse the trend in the amount of bilingual teaching experience noted in late-exit grades three and four by reporting eight or more years of such teaching experience.

Table 61

Ethnic Background of Teachers, by Program and Grade

Grade	Ethnic Group	Program Immersion %	 Early-Exit %	 Late-Exit %
K	American Indian or Alaskan Native	0.0	0.0	0.0
	Asian or Pacific Islander	0.0	0.0	0.0
	Filipino	0.0	3.5	0.0
	Hispanic or Spanish	44.3	48.5	89.1
	Black not of Hispanic Origin	0.0	0.0	3.0
	White not of Hispanic Origin	55.7	48.0	8.0
1	American Indian or Alaskan Native	0.0	0.0	0.0
	Asian or Pacific Islander	0.0	0.0	0.0
	Filipino	2.2	1.1	0.0
	Hispanic or Spanish	67.5	52.0	93.0
	Black not of Hispanic Origin	5.5	0.0	4.2
	White not of Hispanic Origin	24.8	46.9	2.8
2	American Indian or Alaskan Native	0.0	7.2	0.0
	Asian or Pacific Islander	0.0	0.0	0.0
	Filipino	0.0	0.0	0.0
	Hispanic or Spanish	60.5	53.3	91.5
	Black not of Hispanic Origin	1.8	5.0	2.3
	White not of Hispanic Origin	37.6	34.5	6.2
3	American Indian or Alaskan Native	0.0	0.0	2.9
	Asian or Pacific Islander	0.0	0.0	2.9
	Filipino	0.0	0.0	0.0
	Hispanic or Spanish	45.3	64.8	78.4
	Black not of Hispanic Origin	8.1	0.0	5.9
	White not of Hispanic Origin	46.6	35.2	9.9

Table 61 (continued)

Grade	Ethnic Group	Program Immersion %	 Early-Exit %	 Late-Exit %
4	American Indian or Alaskan Native	0.0	0.0	0.0
	Asian or Pacific Islander	0.0	0.0	0.0
	Filipino	0.0	0.0	0.0
	Hispanic or Spanish	18.1	42.0	86.5
	Black not of Hispanic Origin	0.0	0.0	0.0
	White not of Hispanic Origin	81.9	58.0	13.5
5	American Indian or Alaskan Native	*	*	0.0
	Asian or Pacific Islander	*	*	0.0
	Filipino	*	*	0.0
	Hispanic or Spanish	*	*	75.4
	Black not of Hispanic Origin	*	*	5.4
	White not of Hispanic Origin	*	*	19.2
6	American Indian or Alaskan Native	*	*	0.0
	Asian or Pacific Islander	*	*	0.0
	Filipino	*	*	0.0
	Hispanic or Spanish	*	*	56.3
	Black not of Hispanic Origin	*	*	18.3
	White not of Hispanic Origin	*	*	25.4
	Weighted Number of responses	181	180	181

ERIC

Table 62

Teachers' Highest Educational Degree, by Program and Grade

Grade	Degree	Program Immersion %	 Early-Exit %	 Late-Exit %
K	Bachelor's	57.7	57.7	29.8
	Master's	42.3	42.3	66.2
	Doctoral	0.0	0.0	0.0
	Other	0.0	0.0	0.0
	No degree above H.S. diploma	0.0	0.0	4.0
1	Bachelor's	71.5	61.8	33.6
	Master's	28.5	35.1	62.4
	Doctoral	0.0	1.0	0.0
	Other	0.0	2.2	4.1
	No degree above H.S. diploma	0.0	0.0	0.0
2	Bachelor's	66.4	67.9	34.8
	Master's	33.6	24.9	58.5
	Doctoral	0.0	2.4	2.3
	Other	0.0	4.8	4.5
	No degree above H.S. diploma	0.0	0.0	0.0
3	Bachelor's	52.6	62.4	47.0
	Master's	42.5	37.6	51.5
	Doctoral	4.8	0.0	0.0
	Other	0.0	0.0	1.5
	No degree above H.S. diploma	0.0	0.0	0.0

Table 62 (continued)

Grade	Degree	Immersion %	Program Early-Exit %	Late-Exit %
4	Bachelor's	100.0	86.8	30.3
	Master's	0.0	13.2	69.7
	Doctoral	0.0	0.0	0.0
	Other	0.0	0.0	0.0
	No degree above H.S. diploma	0.0	0.0	0.0
5	Bachelor's	*	*	10.8
	Master's	*	*	89.2
	Doctoral	*	*	0.0
	Other	*	*	0.0
	No degree above H.S. diploma	*	*	0.0
6	Bachelor's	*	*	28.2
	Master's	*	*	71.8
	Doctoral	*	*	0.0
	Other	*	*	0.0
	No degree above H.S. diploma	*	*	0.0
	Weighted Number of responses	181	181	181

ERIC

Table 63

Percentage of Teachers with Bilingual and ESL Certification, by Program and Grade

Grade	Type of Certification	Immersion %	Program Early-Exit %	Late-Exit %
K	Bilingual	65.0	82.5	97.0
	ESL	34.2	34.2	53.8
1	Bilingual	68.3	73.1	91.6
	ESL	41.6	41.9	15.4
2	Bilingual	61.7	56.0	84.2
	ESL	14.8	34.8	41.0
3	Bilingual	54.2	52.8	82.4
	ESL	18.3	37.8	38.4
4	Bilingual	18.1	19.8	78.9
	ESL	36.2	20.6	24.8
5	Bilingual	*	*	91.9
	ESL	*	*	57.3
6	Bilingual	*	*	84.5
	ESL	*	*	57.0
	Weighted Number of responses	181,180	180,180	181,180

Table 64

Percentage of Teachers in Each Program
Reporting Special Courses or In-Service Training, by Grade

Grade	Area of Training	Program Immersion %	 Early-Exit %	 Late-Exit %
K	ESL	75.6	84.3	91.3
	Second language acquisition	85.8	80.8	79.8
	First language development	75.0	63.3	65.1
	Multicultural education	77.4	80.7	73.9
	Bilingual education	90.0	91.1	100.0
	Language assessment	63.0	75.6	73.9
1	ESL	87.3	83.6	85.5
	Second language acquisition	85.9	83.6	79.3
	First language development	71.9	80.4	72.4
	Multicultural education	79.7	89.9	54.0
	Bilingual education	88.7	90.1	93.7
	Language assessment	78.1	67.7	67.6
2	ESL	80.0	74.2	75.3
	Second language acquisition	72.5	74.1	88.1
	First language development	65.2	58.2	83.4
	Multicultural education	85.2	66.8	74.1
	Bilingual education	88.1	80.5	97.6
	Language assessment	58.2	63.2	81.0
3	ESL	87.5	74.3	89.9
	Second language acquisition	82.4	74.4	81.1
	First language development	63.9	74.4	68.9
	Multicultural education	88.5	76.4	80.1
	Bilingual education	91.1	78.5	91.3
	Language assessment	77.6	70.6	68.6

Table 64 (continued)

Grade	Area of Training	Program Immersion %	Early-Exit %	Late-Exit %
4	ESL	77.1	71.2	69.2
	Second language acquisition	41.0	50.4	70.8
	First language development	18.1	43.5	65.8
	Multicultural education	59.0	51.2	73.4
	Bilingual education	41.0	56.2	81.1
	Language assessment	22.9	36.9	71.0
5	ESL	*	*	81.7
	Second language acquisition	*	*	81.7
	First language development	*	*	67.5
	Multicultural education	*	*	72.9
	Bilingual education	*	*	86.9
	Language assessment	*	*	71.1
6	ESL	*	*	100.0
	Second language acquisition	*	*	95.3
	First language development	*	*	100.0
	Multicultural education	*	*	95.3
	Bilingual education	*	*	100.0
	Language assessment	*	*	100.0
Range of Weighted number of responses		180 - 181	178 - 179	178 - 181

ERIC

Table 65

Mean Number of Years of Teaching Experience, by Program and Grade

Grade		Program: Immersion	Program: Early-Exit	Program: Late-Exit
K	General Teaching Experience	10.7	9.6	11.5
	Bilingual Teaching Experience	5.0	6.7	7.2
	ESL Teaching Experience	3.1	0.8	0.5
1	General Teaching Experience	10.5	9.7	11.6
	Bilingual Teaching Experience	6.4	6.3	8.2
	ESL Teaching Experience	1.5	1.6	0.3
2	General Teaching Experience	10.8	11.4	10.4
	Bilingual Teaching Experience	5.9	6.4	7.7
	ESL Teaching Experience	1.4	2.0	2.4
3	General Teaching Experience	15.2	10.4	7.7
	Bilingual Teaching Experience	6.8	5.1	6.6
	ESL Teaching Experience	3.2	2.3	0.2
4	General Teaching Experience	10.9	7.0	9.0
	Bilingual Teaching Experience	6.4	3.0	5.9
	ESL Teaching Experience	0.0	0.0	3.0
5	General Teaching Experience	*	*	10.9
	Bilingual Teaching Experience	*	*	8.2
	ESL Teaching Experience	*	*	1.8
6	General Teaching Experience	*	*	10.8
	Bilingual Teaching Experience	*	*	9.0
	ESL Teaching Experience	*	*	2.7
	Weighted Number of responses	181,165,168	181,167,168	181,159,163

Language abilities. All of the teachers in the study report that they speak English (see Table 66). There are grade level differences within and between programs regarding the proportion of teachers who have English as their first or second language. Consistently across grade levels, proportionately fewer late-exit teachers have English as their first language (17.5% to 52.0%) as compared to either immersion strategy (53.8% to 100.0%) or early-exit (60.8% to 92.6%) teachers. Conversely, more late-exit teachers (48.0% to 85.0%) have Spanish as their first language as compared to either immersion strategy (0.0% to 49.2%) or early-exit (7.4% to 42.2%) teachers. These teacher self-reports are consistent with the results of the individualized assessment of teacher English and Spanish oral language proficiency reported earlier. Of concern from the perspective of program implementation is the proportion of immersion strategy (18.6% to 59.0%) and early-exit (14.8% to 28.0%) teachers reporting that they do not speak Spanish. As noted earlier, while the Spanish oral language proficiency of immersion strategy teachers may not be sufficient to teach in it effectively, their Spanish receptive skills still might be sufficiently high to understand and assess what their students are doing. In contrast, the concern remains for early-exit teachers, whose assessed and self-reported skills raise the issue of the quality of program implementation, as, to some extent, they must teach in Spanish.

Table 66

Percentage of Teachers Who Speak English and Spanish, by Program and Grade

		Program								
		Immersion			Early-Exit			Late-Exit		
Grade	Language	First Lang. %	Second Lang. %	Do Not Speak %	First Lang. %	Second Lang. %	Do Not Speak %	First Lang. %	Second Lang. %	Do Not Speak %
K	English	85.0	15.0	0.0	60.8	39.2	0.0	26.9	73.1	0.0
	Spanish	15.0	59.4	25.6	42.2	40.3	17.5	73.1	23.9	3.0
1	English	53.8	46.2	0.0	74.6	25.5	0.0	17.5	82.5	0.0
	Spanish	49.2	23.8	27.0	26.5	53.9	19.7	85.0	12.8	2.2
2	English	59.8	40.2	0.0	74.8	25.2	0.0	20.8	79.2	0.0
	Spanish	40.2	31.4	28.5	26.1	52.6	21.3	79.2	18.5	2.3
3	English	84.6	15.4	0.0	67.0	33.0	0.0	34.5	65.5	0.0
	Spanish	20.4	61.1	18.6	29.7	42.4	28.0	65.5	28.6	5.9
4	English	100.0	0.0	0.0	92.6	7.4	0.0	27.8	72.2	0.0
	Spanish	0.0	41.0	59.0	7.4	77.8	14.8	67.3	22.7	10.0
5	English	*	*	*	*	*	*	20.3	79.7	0.0
	Spanish	*	*	*	*	*	*	76.6	23.4	0.0
6	English	*	*	*	*	*	*	52.0	48.0	0.0
	Spanish	*	*	*	*	*	*	48.0	52.0	0.0
	Weighted Number of responses		180,180			180,180			180,180	

ERIC

<u>Are classrooms comparable between programs with respect to extra help and student behavior?</u>

Across grades within and between programs, approximately two-thirds or more of the teachers are in self-contained classrooms (Table 67). With minor exceptions, the remaining teachers are in team-teaching positions. The services of other resource teachers or instructional aides is negligible or non-existent in each program (see Table 68). Teachers report an average work day of 6.3 to 7.5 hours (see Table 69). Of this, 4.5 to 6.0 hours are spent on instruction. With the exception of kindergarten immersion strategy and early-exit classrooms, there is little grade level variation within programs regarding the amount of instructional or total work time. There are no major program differences.

With one exception, across grades within and between programs, the majority of teachers report that student misbehavior interferes with instruction to a small extent or not at all (see Table 70). The exception is in fourth grade immersion strategy classrooms where 77.2% of the teachers indicate that student misbehavior interferes with instruction to a moderate or great extent. Notwithstanding, there is a trend in each program of more teachers reporting that student misbehavior interferes with instruction to a moderate extent as grade level increases.

<u>What is the language proficiency level of target students' classmates? Does this differ by program?</u>

Across grades within programs, with the exception of fourth grade immersion strategy teachers, there is a trend towards a decrease in the proportion of teachers reporting that "some language minority students are proficient in English, but most have limited proficiency" and an increase in the proportion of those reporting that "most language minority students are proficient in English" (see Table 71). This change presumably reflects the increased English language skills among language minority children in each program. Interestingly, in the

ERIC

immersion strategy and early-exit programs, there is a marked increase both in the proportion of teachers who report students as being "quite limited" and in those reporting that "most language minority students are proficient in English" as one moves from kindergarten to first grade. These changes may reflect increased expectations of first grade teachers regarding English proficiency, as well as progress in developing English language skills by language minority students from kindergarten to first grade. While the proportion of teachers judging students as being "quite limited" dramatically drops in immersion strategy programs from first to second grade and totally disappears in grade three, the proportion of this group increases in both the early-exit and late-exit programs from grades one to two. It is not until grade three that the proportion of teachers judging children as being "quite limited" begins to decrease in both early-exit and late-exit classrooms. This decline continues until the proportion of teachers reporting students as being "quite limited English proficient" disappears in fourth grade early-exit classrooms and in sixth grade late-exit classrooms. The slower rate at which language minority students are judged by their teachers as showing an increase in their English language skills in the early- and late-exit program presumably reflects the instructional objectives of each. That is, both bilingual programs exhibit a slower pace of acquisition of English language skills, with the late-exit program allowing for the most flexibility. Their approaches are reflected in their judgments as to the relative English language proficiency of their students.

Table 67

Percentage of Teachers in Self-Contained Classrooms and Team Teaching Situations, by Program and Grade

Grade		Program: Immersion %	Program: Early-Exit %	Program: Late-Exit %
K	Self-Contained	71.2	65.9	81.4
	Team-teaching	28.9	34.1	18.6
	Other	0.0	0.0	0.0
1	Self-Contained	96.3	72.4	68.4
	Team-teaching	3.7	22.1	29.5
	Other	0.0	5.5	2.1
2	Self-Contained	87.4	73.9	65.5
	Team-teaching	12.6	26.2	34.5
	Other	0.0	0.0	0.0
3	Self-Contained	82.7	80.2	73.3
	Team-teaching	7.6	19.8	26.7
	Other	9.7	0.0	0.0
4	Self-Contained	63.9	77.8	76.1
	Team-teaching	36.2	22.2	21.3
	Other	0.0	0.0	2.6
5	Self-Contained	*	*	78.6
	Team-teaching	*	*	21.4
	Other	*	*	0.0
6	Self-Contained	*	*	65.1
	Team-teaching	*	*	34.9
	Other	*	*	0.0
	Weighted Number of responses	181	181	181

Table 68

Mean Number of Hours of Services Provided by Classroom Aides and Resource Teachers, by Program and Grade

Grade		Immersion	Program Early-Exit	Late-Exit
K	Classroom Aide	1.7	3.3	1.4
	Resource Teacher	0.4	0.8	0.8
1	Classroom Aide	2.1	1.9	2.0
	Resource Teacher	1.3	0.6	0.8
2	Classroom Aide	2.9	1.3	1.1
	Resource Teacher	2.0	1.4	1.0
3	Classroom Aide	1.8	1.4	0.6
	Resource Teacher	1.1	1.2	1.6
4	Classroom Aide	1.4	0.4	0.5
	Resource Teacher	0.7	1.5	1.9
5	Classroom Aide	*	*	0.7
	Resource Teacher	*	*	0.8
6	Classroom Aide	*	*	0.3
	Resource Teacher	*	*	0.8
	Weighted Number of responses	180,181	180,181	180,181

Table 69

Mean Number of Hours of Work and Instruction, by Program and Grade

Grade		Program Immersion	Early-Exit	Late-Exit
K	Instructional Time	4.5	4.5	5.0
	Total Work Time	6.3	6.4	6.4
1	Instructional Time	5.6	5.6	5.2
	Total Work Time	7.0	6.9	6.3
2	Instructional Time	5.5	5.5	5.3
	Total Work Time	6.9	6.8	6.4
3	Instructional Time	5.3	5.4	5.4
	Total Work Time	6.9	6.9	6.6
4	Instructional Time	5.8	6.0	5.0
	Total Work Time	7.4	7.5	6.4
5	Instructional Time	*	*	5.0
	Total Work Time	*	*	6.5
6	Instructional Time	*	*	5.2
	Total Work Time	*	*	6.6
	Weighted Number of responses	179,180	176,180	176,180

ERIC

Table 70

Extent to Which Teachers Report Misbehavior Interferes With Instruction, by Program and Grade

Grade		Immersion %	Program Early-Exit %	Late-Exit %
K	Great extent	3.2	24.9	0.0
	Moderate extent	6.3	17.6	10.0
	Small extent	60.8	47.3	62.2
	Not at all	29.7	10.3	27.8
1	Great extent	3.4	15.5	11.3
	Moderate extent	24.0	16.3	11.3
	Small extent	54.8	44.1	66.7
	Not at all	17.9	24.0	10.7
2	Great extent	5.7	3.6	10.6
	Moderate extent	15.4	22.7	15.3
	Small extent	46.2	38.5	55.4
	Not at all	32.6	35.3	18.7
3	Great extent	9.7	3.9	10.9
	Moderate extent	17.8	31.2	23.3
	Small extent	62.9	47.3	46.6
	Not at all	9.7	17.6	19.2
4	Great extent	41.0	7.4	4.9
	Moderate extent	36.2	20.6	34.6
	Small extent	0.0	64.6	40.7
	Not at all	22.9	7.4	19.7

Table 70 (continued)

Grade		Immersion %	Program Early-Exit %	Late-Exit %
5	Great extent	*	*	7.9
	Moderate extent	*	*	13.5
	Small extent	*	*	56.7
	Not at all	*	*	21.9
6	Great extent	*	*	0.0
	Moderate extent	*	*	22.2
	Small extent	*	*	48.4
	Not at all	*	*	29.4
	Weighted number of responses	178	181	175

ERIC

Table 71

Teachers' Descriptions of Students' English Language Proficiency, by Program and Grade

Grade		Program Immersion %	 Early-Exit %	 Late-Exit %
K	All LM students are quite limited in English	7.5	7.7	29.2
	Some LM students are proficient in English, but most have limited proficiency	81.9	92.3	70.8
	Most LM students are proficient in English	10.6	0.0	0.0
1	All LM students are quite limited in English	15.7	22.0	22.8
	Some LM students are proficient in English, but most have limited proficiency	57.2	64.7	73.0
	Most LM students are proficient in English	27.1	13.3	4.2
2	All LM students are quite limited in English	5.6	25.4	32.9
	Some LM students are proficient in English, but most have limited proficiency	58.7	46.1	56.1
	Most LM students are proficient in English	35.7	28.5	11.0
3	All LM students are quite limited in English	0.0	17.8	7.7
	Some LM students are proficient in English, but most have limited proficiency	59.3	50.6	58.2
	Most LM students are proficient in English	40.7	31.6	34.1
4	All LM students are quite limited in English	0.0	0.0	7.5
	Some LM students are proficient in English, but most have limited proficiency	81.9	7.4	42.0
	Most LM students are proficient in English	18.1	92.6	50.6
5	All LM students are quite limited in English	*	*	10.4
	Some LM students are proficient in English, but most have limited proficiency	*	*	42.0
	Most LM students are proficient in English	*	*	47.6
6	All LM students are quite limited in English	*	*	0.0
	Some LM students are proficient in English, but most have limited proficiency	*	*	12.7
	Most LM students are proficient in English	*	*	87.3
	Weighted Number of responses	181	178	180

ERIC

When Spanish language skills are considered, there are marked differences by grade level within and between programs (see Table 72). Noteworthy is that at kindergarten, a greater proportion of late-exit teachers judge their students as being highly proficient in Spanish (54.7%) than either immersion strategy (19.6%) or early-exit (1.9%) teachers. From a slightly different perspective, greater proportions of immersion strategy and early-exit teachers indicate that their language minority students represent a mixed range of Spanish language skills (80.4% and 98.1%, respectively) than late-exit teachers (45.3%). Interestingly, the proportion of immersion strategy teachers judging their students as being highly proficient in Spanish dramatically increases from kindergarten to first grade, subsequently decreases in second and third grades, and disappears in grade four. A similar pattern is exhibited by early-exit teachers. This change in teacher judgments over time is not expected, to say the least, especially within the immersion strategy program. There is no logical explanation, from the students' point of view, as to why there should be a sudden increase in their Spanish language proficiency from kindergarten to first grade in programs which do not provide for Spanish language instruction, or do so in a limited way. One can only surmise that either immersion strategy and early-exit kindergarten teachers do not have the necessary skills to assess the language proficiency of their students, and/or that they view their students' Spanish language proficiency differently than do teachers in the upper grades. As we did not observe any major grade level differences with respect to special teaching credentials or coursework, in all likelihood teachers in the upper grades may look at Spanish language proficiency as a reason why some of their students may not be progressing in their English language skills as desired. In marked contrast, approximately half of the late-exit teachers in kindergarten through fourth grade consistently judge their students as being highly proficient in Spanish.

Within each program, as grade level increases, the proportion of teachers judging language minority students to be highly proficient in

Spanish decreases as grade level increases, albeit the rate at which this decrease in Spanish language proficiency occurs is different between programs. By grade three only a few immersion strategy (8.7%) and early-exit (4.1%) teachers judge their students as being highly proficient in Spanish. That is, the majority of these teachers do not judge their students as being highly proficient in Spanish. This decrease indicates the rapid loss of Spanish language skills among language minority students, which in turn reflects the long-term program objectives of both immersion strategy and early-exit programs. In contrast, it is not until grades five and six that about one-quarter and one-third, respectively, of the late-exit teachers judge their students as being highly proficient in Spanish. Prior to this, 41.4% to 54.7% of the teachers rate their students as being highly proficient in Spanish. These teacher judgments reflect the late-exit program objective of maintaining a student's primary language skills.

<u>When do project teachers begin language instruction in English and Spanish, and does this differ by program?</u>

There are marked grade level differences within and between programs as to when teachers begin language instruction in English and Spanish. With the exception of kindergarten, almost all immersion strategy teachers report providing instruction in English reading (see Table 73). In kindergarten, slightly more than three-fourths of the teachers report providing English reading instruction. There is a general, although uneven, increase in the proportion of early-exit teachers providing instruction in English reading as grade level increases (41.7% to 92.6%). Among the late-exit teachers, there is a steady increase in the proportion of teachers providing instruction in English reading, from 19.7% in kindergarten to 42.0% in the second grade. This proportion almost doubles in grade three (83.4%), and remains at this high level through grade six. These grade level differences within and between programs are consistent with the instructional model of each program.

Table 72

Teachers' Descriptions of Students' Spanish Language Proficiency, by Program and Grade

Grade		Immersion %	Program Early-Exit %	Late-Exit %
K	All LM students are highly proficient in Spanish	19.6	1.9	54.7
	Some LM students are proficient in Spanish and some are not	80.4	98.1	45.3
	No LM students are proficient in Spanish	0.0	0.0	0.0
1	All LM students are highly proficient in Spanish	32.2	13.4	52.0
	Some LM students are proficient in Spanish and some are not	66.2	85.6	48.0
	No LM students are proficient in Spanish	1.7	1.0	0.0
2	All LM students are highly proficient in Spanish	24.2	27.0	43.3
	Some LM students are proficient in Spanish and some are not	75.8	71.7	48.9
	No LM students are proficient in Spanish	0.0	1.4	7.7
3	All LM students are highly proficient in Spanish	8.7	4.1	41.4
	Some LM students are proficient in Spanish and some are not	91.3	95.9	58.6
	No LM students are proficient in Spanish	0.0	0.0	0.0
4	All LM students are highly proficient in Spanish	0.0	0.0	51.7
	Some LM students are proficient in Spanish and some are not	100.0	77.8	48.3
	No LM students are proficient in Spanish	0.0	22.2	0.0
5	All LM students are highly proficient in Spanish	*	*	26.4
	Some LM students are proficient in Spanish and some are not	*	*	68.4
	No LM students are proficient in Spanish	*	*	5.2
6	All LM students are highly proficient in Spanish	*	*	32.5
	Some LM students are proficient in Spanish and some are not	*	*	58.3
	No LM students are proficient in Spanish	*	*	9.2
	Weighted Number of responses	167	171	178

ERIC

Table 73

Percentage of Teachers Providing Formal Instruction in English Reading to LEP Children, by Program and Grade

Grade	Program Immersion %	 Early-Exit %	 Late-Exit %
K	77.5	41.7	19.7
1	100.0	63.9	32.2
2	98.3	71.2	42.0
3	96.2	64.5	83.4
4	100.0	92.6	89.8
5	*	*	89.2
6	*	*	81.1
Weighted Number of responses	181	180	181

A basic tenet, if not the central hypothesis, of the bilingual programs is that limited English proficient students can acquire their English language skills more rapidly if their primary language skills are developed to a critical threshold level. With this in mind, it is important to determine the extent to which teachers in each of the three programs use a student's proficiency in Spanish as a criterion for beginning instruction in English reading. Grade level differences within and between programs are noted. Within the immersion strategy program, approximately one-fifth or less of the teachers use their students' Spanish language proficiency as a criterion for determining when to begin instruction in English reading (see Table 74). Moreover, there is no consistent pattern across grades with which immersion strategy teachers do so. When asked to identify when they begin formal instruction in English reading with limited English proficient students, at least half or more of immersion strategy teachers in three of the five grades indicate that they begin English reading instruction immediately, without regard to Spanish language proficiency (see Table 75). Alternately, at least half of the teachers in the remaining two

ERIC

grades note that they begin English reading instruction only after Spanish literacy is developed.

Among the early-exit teachers there is a consistent decrease in the proportion of teachers using Spanish proficiency as a criterion for introducing English reading instruction, although the proportion remains larger than half through third grade (see Table 74). This decrease probably reflects the fact that as students are introduced to English reading, there is no need for teachers in later grades to consider a student's Spanish language proficiency. Teachers would automatically continue with English reading instruction. Clearly, the majority of early-exit teachers through grade three indicate that they begin instruction in English reading with limited English proficient students only after Spanish literacy is developed (see Table 75). By grade four, early-exit teachers report using other criteria to determine when to begin formal instruction in English reading.

In contrast to immersion strategy teachers, yet with some similarity to early-exit teachers, the majority of late-exit teachers consistently report using the child's Spanish language proficiency to determine readiness for English reading instruction (see Table 74). In kindergarten, only half of the teachers report that they do so, apparently reflecting the fact that many kindergarten late-exit programs do not begin formal instruction in English reading until first grade. At that time they use the child's Spanish language proficiency to determine readiness for English reading instruction. More specifically, across grades, the majority of late-exit teachers begin formal instruction in English reading with limited English proficient students only after Spanish literacy is developed (see Table 75).

ERIC

Table 74

Percentage of Teachers Using Proficiency in Spanish to Determine Readiness for English Reading Instruction, by Program and Grade

Grade	Program Immersion %	 Early-Exit %	 Late-Exit %
K	16.0	70.3	50.0
1	4.6	66.5	85.5
2	22.5	61.8	88.6
3	13.7	55.0	82.2
4	22.9	0.0	74.3
5	*	*	91.9
6	*	*	65.6
Weighted Number of responses	146	147	148

ERIC

Table 75

Percentage of Teachers Beginning Formal English Reading Instruction with LEP Students, by Program and Grade

Grade	Time	Immersion %	Program Early-Exit %	Late-Exit %
K	After Spanish literacy is developed	50.0	100.0	100.0
	At same time Spanish literacy is developed	0.0	0.0	0.0
	Immediately, no Spanish literacy	0.0	0.0	0.0
	Other	50.0	0.0	0.0
1	After Spanish literacy is developed	0.0	100.0	92.8
	At same time Spanish literacy is developed	0.0	0.0	7.2
	Immediately, no Spanish literacy	100.0	0.0	0.0
	Other	0.0	0.0	0.0
2	After Spanish literacy is developed	25.4	81.3	97.5
	At same time Spanish literacy is developed	0.0	9.5	2.5
	Immediately, no Spanish literacy	58.4	1.7	0.0
	Other	16.2	7.5	0.0
3	After Spanish literacy is developed	74.0	85.6	84.5
	At same time Spanish literacy is developed	0.0	14.4	9.5
	Immediately, no Spanish literacy	26.0	0.0	0.0
	Other	0.0	0.0	6.1
4	After Spanish literacy is developed	0.0	0.0	73.1
	At same time Spanish literacy is developed	0.0	0.0	16.1
	Immediately, no Spanish literacy	100.0	0.0	6.4
	Other	0.0	100.0	4.4

Table 75 (continued)

Grade	Time	Immersion %	Program Early-Exit %	Late-Exit %
5	After Spanish literacy is developed	*	*	63.1
	At same time Spanish literacy is developed	*	*	36.9
	Immediately, no Spanish literacy	*	*	0.0
	Other	*	*	0.0
6	After Spanish literacy is developed	*	*	58.1
	At same time Spanish literacy is developed	*	*	41.9
	Immediately, no Spanish literacy	*	*	0.0
	Other	*	*	0.0

Data for this item should be interpreted with extreme caution.

	Immersion	Early-Exit	Late-Exit
Weighted Number of responses	25	87	122

ERIC

When formal instruction in Spanish reading is considered, the majority of immersion strategy teachers do not teach Spanish reading (see Table 76). While approximately two-thirds of the kindergarten and first grade early-exit teachers report beginning formal instruction in Spanish reading "immediately, before English literacy," this proportion decreases steadily to 44.1% in grade three. By grade four early-exit teachers do not provide formal instruction in Spanish reading. A large proportion of late-exit teachers in kindergarten through third grade report providing formal instruction in Spanish reading prior to English literacy (62.6% to 90.0%). In the fourth and fifth grades slightly less than half (46.0% and 48.1%, respectively) provide such instruction. Approximately one-fifth (21.9%) of the sixth grade teachers provide formal instruction in Spanish reading skills prior to English literacy. However, at least two-thirds of the late-exit teachers at each grade report providing formal Spanish reading instruction, regardless of students' English literacy.

The patterns reported are consistent with the immersion strategy and early-exit instructional models. Clearly, the fact that large proportions of late-exit teachers in the upper grades (up to one-third) do not provide Spanish reading instruction is of concern with respect to fidelity of treatment. This late-exit pattern is consistent with the data from classroom observations indicating a decline in the use of the student's primary language in the classroom by late-exit teachers in the upper grades. (However, once again this pattern is attributed to one of the three late-exit sites.)

Table 76

Percentage of Teachers Beginning Formal Spanish Reading Instruction with LEP Students, by Program and Grade

Grade	Time	Program: Immersion %	Early-Exit %	Late-Exit %
K	Immediately, before English literacy	0.0	63.3	79.6
	At same time as English literacy	0.0	3.3	5.9
	After English literacy is introduced	0.0	0.0	0.0
	Do not teach Spanish literacy	52.0	20.7	8.7
	Not taught in this program	48.1	12.7	5.8
	Other	0.0	0.0	0.0
1	Immediately, before English literacy	0.0	66.5	90.0
	At same time as English literacy	2.4	7.9	6.9
	After English literacy is introduced	7.1	0.0	0.0
	Do not teach Spanish literacy	32.0	12.8	3.1
	Not taught in this program	57.1	8.1	0.0
	Other	1.5	4.6	0.0
2	Immediately, before English literacy	3.8	55.7	75.7
	At same time as English literacy	0.0	3.8	12.5
	After English literacy is introduced	0.0	0.0	0.0
	Do not teach Spanish literacy	24.2	24.1	8.5
	Not taught in this program	72.0	11.3	0.0
	Other	0.0	5.1	3.4
3	Immediately, before English literacy	3.8	44.1	62.6
	At same time as English literacy	0.0	0.0	11.7
	After English literacy is introduced	0.0	4.1	0.0
	Do not teach Spanish literacy	41.0	36.0	19.8
	Not taught in this program	55.2	11.7	5.9
	Other	0.0	4.1	0.0

Table 76 (continued)

Grade	Time	Program Immersion %	Early-Exit %	Late-Exit %
4	Immediately, before English literacy	0.0	0.0	46.0
	At same time as English literacy	0.0	0.0	22.1
	After English literacy is introduced	0.0	0.0	5.5
	Do not teach Spanish literacy	45.8	92.6	22.6
	Not taught in this program	54.2	0.0	3.8
	Other	0.0	7.4	0.0
5	Immediately, before English literacy	*	*	48.1
	At same time as English literacy	*	*	41.4
	After English literacy is introduced	*	*	0.0
	Do not teach Spanish literacy	*	*	6.6
	Not taught in this program	*	*	3.8
	Other	*	*	0.0
6	Immediately, before English literacy	*	*	21.9
	At same time as English literacy	*	*	38.9
	After English literacy is introduced	*	*	6.4
	Do not teach Spanish literacy	*	*	32.9
	Not taught in this program	*	*	0.0
	Other	*	*	0.0
	Weighted Number of responses	172	173	172

ERIC

In what language do project teachers teach? Does this differ by program?

Teachers in each of the programs on the average report using English for instruction in a manner consistent with their respective instructional models. Immersion strategy teachers across grades consistently report that they use only English for instruction or that they use English supplemented with Spanish when teaching math, science, social studies, and ethnic heritage (61.5% to 100.0%) (see Table 77). More than two-thirds of early-exit teachers report using some Spanish in kindergarten through grade three for math, science and social studies (i.e., primary only, primary supplemented with English, and/or equal primary and English) (see Table 78), while more than half report using no Spanish in grade four. There is a definite decline in the use of Spanish and an increase in the use of English (i.e., English only; English with primary supplement) as grade level increases. Noticeably different, late-exit teachers report the greatest use of Spanish (i.e., primary only; primary with English supplement) from kindergarten through sixth grade (see Table 79). The exclusive or near exclusive use of Spanish decreases as grade level increases, changing to a more equal use of Spanish and English. Nonetheless, it is not until grade six that late-exit teachers report shifting to greater use of English than Spanish (i.e., English with primary supplement and English only). These teacher self-reported patterns of language use are consistent with the instructional model of each program.

ERIC

Table 77

Percentage of Immersion Strategy Program Teachers Reporting Classroom Language Use, by Grade

		Subject			
Grade		Math %	Science %	Social Studies %	Ethnic Heritage %
K	Primary only	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	3.7
	Equal primary and English	0.0	0.0	0.0	6.9
	English w/primary to supplement	59.4	52.5	58.8	58.8
	English only	40.6	43.7	37.5	27.6
	Not taught	0.0	3.7	3.7	3.0
1	Primary only	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	1.8	1.8	9.5
	Equal primary and English	1.4	3.3	1.4	5.9
	English w/primary to supplement	72.1	64.8	66.6	50.8
	English only	26.4	30.1	30.1	20.1
	Not taught	0.0	0.0	0.0	13.8
2	Primary only	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	2.2
	Equal primary and English	0.0	4.4	4.4	11.1
	English w/primary to supplement	44.4	48.9	46.2	37.8
	English only	55.6	44.9	47.6	41.0
	Not taught	0.0	1.8	1.8	7.9
3	Primary only	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	0.0
	Equal primary and English	5.9	6.2	6.2	11.2
	English w/primary to supplement	53.4	41.7	37.7	23.5
	English only	40.7	47.1	47.1	38.0
	Not taught	0.0	5.1	9.1	27.3

ERIC

Table 77 (continued)

Grade		Subject Math %	 Science %	 Social Studies %	 Ethnic Heritage %
4	Primary only	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	0.0
	Equal primary and English	0.0	0.0	0.0	0.0
	English w/primary to supplement	0.0	36.2	18.1	0.0
	English only	81.9	63.9	81.9	81.9
	Not taught	18.1	0.0	0.0	18.1
Weighted Number of Responses		180	179	179	179

Table 78

Percentage of Early-Exit Program Teachers Reporting Classroom Language Use, by Grade

Grade		Math %	Science %	Social Studies %	Ethnic Heritage %
		Subject			
K	Primary only	29.2	3.7	3.7	7.4
	Primary w/English to supplement	12.4	5.4	5.4	7.0
	Equal primary and English	31.4	32.5	36.0	44.6
	English w/primary to supplement	18.3	46.1	42.5	25.2
	English only	8.8	12.3	12.3	10.7
	Not taught	0.0	0.0	0.0	5.1
1	Primary only	6.7	8.8	10.9	15.4
	Primary w/English to supplement	18.5	5.3	7.5	7.5
	Equal primary and English	10.9	12.3	14.5	12.2
	English w/primary to supplement	48.5	48.4	47.3	33.1
	English only	13.2	18.8	15.5	13.3
	Not taught	2.3	6.5	4.4	18.6
2	Primary only	4.9	9.6	7.2	12.0
	Primary w/English to supplement	12.1	12.0	12.0	7.3
	Equal primary and English	9.5	7.2	9.5	4.8
	English w/primary to supplement	47.1	37.8	40.2	32.8
	English only	26.4	28.6	23.9	20.3
	Not taught	0.0	4.8	7.2	12.8
3	Primary only	4.6	13.7	9.2	13.5
	Primary w/English to supplement	22.5	9.1	17.9	18.1
	Equal primary and English	6.5	10.9	10.9	4.3
	English w/primary to supplement	41.6	35.0	28.3	28.3
	English only	24.9	31.4	29.1	17.9
	Not taught	0.0	0.0	4.6	17.9

ERIC

Table 78 (continued)

Grade		Subject Math %	Science %	Social Studies %	Ethnic Heritage %
4	Primary only	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	0.0
	Equal primary and English	14.2	0.0	0.0	0.0
	English w/primary to supplement	27.5	41.7	41.7	27.5
	English only	58.3	58.3	58.3	6.6
	Not taught	0.0	0.0	0.0	65.9
Weighted Number of Responses		177	177	177	177

Table 79

Percentage of Late-Exit Program Teachers Reporting Classroom Language Use, by Grade

Grade		Math %	Science %	Social Studies %	Ethnic Heritage %
		Subject			
K	Primary only	60.1	58.5	60.9	69.9
	Primary w/English to supplement	27.0	34.1	32.2	19.6
	Equal primary and English	7.0	0.0	0.0	4.2
	English w/primary to supplement	3.0	4.2	4.0	0.0
	English only	0.0	3.2	3.0	0.0
	Not taught	3.0	0.0	0.0	6.3
1	Primary only	53.6	39.7	43.8	43.1
	Primary w/English to supplement	29.2	25.0	20.8	20.4
	Equal primary and English	7.2	2.9	5.8	13.4
	English w/primary to supplement	10.1	26.0	23.1	12.1
	English only	0.0	0.0	0.0	0.0
	Not taught	0.0	6.5	6.5	11.1
2	Primary only	33.7	47.7	63.5	57.9
	Primary w/English to supplement	44.2	9.3	14.0	7.1
	Equal primary and English	10.2	14.2	12.7	11.2
	English w/primary to supplement	3.2	25.5	9.7	9.7
	English only	0.0	0.0	0.0	0.0
	Not taught	8.7	3.2	0.0	14.2
3	Primary only	6.1	17.3	22.9	34.1
	Primary w/English to supplement	46.9	30.4	32.6	23.4
	Equal primary and English	11.9	8.5	8.2	9.4
	English w/primary to supplement	21.0	21.0	20.6	10.6
	English only	7.9	19.6	11.1	8.4
	Not taught	6.3	3.3	4.7	14.1

Table 79 (continued)

Grade		Subject: Math %	Science %	Social Studies %	Ethnic Heritage %
4	Primary only	0.0	5.3	7.4	21.4
	Primary w/English to supplement	40.0	37.4	32.1	46.5
	Equal primary and English	41.1	23.4	35.8	9.1
	English w/primary to supplement	16.2	20.2	16.5	11.1
	English only	0.0	8.1	2.8	2.8
	Not taught	2.7	5.5	5.5	9.2
5	Primary only	5.5	23.1	17.6	39.6
	Primary w/English to supplement	31.3	33.0	33.0	27.5
	Equal primary and English	44.2	13.8	28.6	21.5
	English w/primary to supplement	11.4	16.9	7.6	7.6
	English only	7.6	13.1	13.1	3.8
	Not taught	0.0	0.0	0.0	0.0
6	Primary only	0.0	7.4	7.4	7.4
	Primary w/English to supplement	18.1	10.7	0.0	21.4
	Equal primary and English	34.3	26.9	44.9	25.5
	English w/primary to supplement	32.9	32.9	25.5	25.5
	English only	7.4	7.4	22.2	14.8
	Not taught	7.4	14.8	0.0	5.5
Weighted Number of Responses		173	169	172	169

Do project teachers assign homework?

Yes. Across grade levels within each program the majority of teachers assign homework two or more times a week (see Table 80). As grade level increases in each program, the proportion of teachers assigning homework more frequently (i.e., ≥ [illegible] times/week) tends to increase. Proportionately more late-exit teachers at each grade level report assigning homework at least three times per week than either immersion strategy or early-exit teachers. Except for second grade, proportionately more late-exit teachers also report grading and returning homework to their students on a frequent basis (i.e., "almost always" and "about 75% of the time") than either immersion strategy or early-exit teachers (see Table 81), although differences between programs are small.

Table 80

Percentage of Teachers Reporting Frequency of Homework Assignments, by Program and Grade

Grade	Frequency	Immersion %	Program Early-Exit %	Late-Exit %
K	Daily	11.6	15.9	50.0
	4 times/week	11.6	0.0	12.7
	3 times/week	30.3	21.2	0.0
	2 times/week	7.7	21.2	8.0
	Weekly	20.0	14.1	15.7
	Less than weekly	6.5	5.3	9.7
	Never	12.4	22.2	4.0
1	Daily	40.7	34.7	59.2
	4 times/week	35.0	46.1	28.1
	3 times/week	15.6	4.5	7.8
	2 times/week	4.3	7.9	4.9
	Weekly	2.2	0.0	0.0
	Less than weekly	0.0	3.4	0.0
	Never	2.2	3.4	0.0
2	Daily	26.9	32.2	62.6
	4 times/week	57.6	38.3	26.8
	3 times/week	11.2	15.5	10.6
	2 times/week	1.7	8.1	0.0
	Weekly	0.0	6.0	0.0
	Less than weekly	2.6	0.0	0.0
	Never	0.0	0.0	0.0
3	Daily	19.4	15.6	56.1
	4 times/week	55.7	50.8	34.1
	3 times/week	3.8	21.6	6.9
	2 times/week	4.8	4.0	2.9
	Weekly	3.8	2.0	0.0
	Less than weekly	0.0	2.0	0.0
	Never	12.5	4.0	0.0

ERIC

Table 80 (continued)

Grade	Frequency	Immersion %	Program Early-Exit %	Late-Exit %
4	Daily	0.0	13.2	58.8
	4 times/week	63.9	35.1	29.5
	3 times/week	0.0	22.2	4.9
	2 times/week	36.2	14.8	3.4
	Weekly	0.0	0.0	0.0
	Less than weekly	0.0	0.0	0.0
	Never	0.0	14.8	3.4
5	Daily	*	*	59.2
	4 times/week	*	*	26.8
	3 times/week	*	*	0.0
	2 times/week	*	*	3.6
	Weekly	*	*	10.4
	Less than weekly	*	*	0.0
	Never	*	*	0.0
6	Daily	*	*	44.9
	4 times/week	*	*	42.4
	3 times/week	*	*	12.7
	2 times/week	*	*	0.0
	Weekly	*	*	0.0
	Less than weekly	*	*	0.0
	Never	*	*	0.0
Weighted	Number of responses	180	181	181

Table 81

Percentage of Teachers Reporting Frequency Homework is Graded and Returned, by Program and Grade

Grade		Program Immersion %	Early-Exit %	Late-Exit %
K	Seldom	9.8	0.0	15.7
	About 25% of the time	9.4	0.0	0.0
	About 50% of the time	0.0	7.0	3.0
	About 75% of the time	0.0	7.3	0.0
	Almost always	68.7	63.5	77.4
	No homework assigned	12.0	22.2	4.0
1	Seldom	1.9	2.2	0.0
	About 25% of the time	1.5	2.3	7.8
	About 50% of the time	6.6	7.8	0.0
	About 75% of the time	21.1	5.5	16.8
	Almost always	66.7	78.9	75.5
	No homework assigned	2.3	3.4	0.0
2	Seldom	0.0	6.0	0.0
	About 25% of the time	2.2	4.7	0.0
	About 50% of the time	8.7	14.4	16.0
	About 75% of the time	26.1	12.0	12.2
	Almost always	63.0	62.9	71.8
	No homework assigned	0.0	0.0	0.0
3	Seldom	0.0	0.0	3.5
	About 25% of the time	4.8	3.9	4.4
	About 50% of the time	4.8	9.6	5.5
	About 75% of the time	4.8	7.9	16.1
	Almost always	73.0	74.7	70.6
	No homework assigned	12.5	4.0	0.0
4	Seldom	0.0	0.0	2.6
	About 25% of the time	18.1	0.0	0.0
	About 50% of the time	22.9	13.2	13.3
	About 75% of the time	0.0	14.2	23.5
	Almost always	59.0	57.8	57.2
	No homework assigned	0.0	14.8	3.4
5	Seldom	*	*	3.6
	About 25% of the time	*	*	0.0
	About 50% of the time	*	*	13.1
	About 75% of the time	*	*	12.4
	Almost always	*	*	70.9
	No homework assigned	*	*	0.0

Table 81 (continued)

Grade		Program Immersion %	Early-Exit %	Late-Exit %
6	Seldom	*	*	0.0
	About 25% of the time	*	*	0.0
	About 50% of the time	*	*	26.6
	About 75% of the time	*	*	6.4
	Almost always	*	*	67.1
	No homework assigned	*	*	0.0
Weighted	Number of responses	180	181	181

<u>How much time do teachers allocate to various content areas</u>?

With one exception, teachers in the three programs allocate approximately the same proportion of instructional time within each subject to math (16.2% to 24.2%), social studies (6.6% to 19.0%), and science (6.5% to 13.1%) (see Table 82). The exception lies in social studies in fourth grade, with late-exit teachers providing approximately half the instructional time (8.7%) that either immersion strategy (19.0%) or early-exit (15.1%) teachers provide. As expected, the three programs differ markedly in the proportion of instructional time allocated to English and Spanish language arts. Kindergarten through third grade immersion strategy teachers report that approximately two-thirds (58.8% to 61.6%) of their instructional time is allocated to English language arts, while fourth grade immersion strategy teachers allocate one-half (51.7%) of their instructional time to English language arts. In contrast, early-exit teachers increase the percentage of English language arts instructional time across grades, from one-third (35.4%) in kindergarten to one-half (50.3%) in fourth grade. Late-exit teachers also report an increase in time allocated to English language arts instruction, although uneven, from less than one-fourth (23.8%) in kindergarten to over one-third (41.7%) in grade four. Conversely, consistent with the instructional model, immersion strategy teachers do not provide instruction in Spanish language arts (0.0 to 1.2%). While early-exit teachers allocate slightly more than one-fourth of their instructional time to Spanish language arts in kindergarten (28.6%), it decreases to less than one-fifth in grade three (17.9%) and is not provided in grade four. Late-exit teachers allocate slightly more than two-fifths (42.7%) of their instructional time to Spanish language arts in kindergarten, slowly decreasing to approximately one-fifth in grade four (21.7%). These patterns are consistent with the instructional models of each program.

Of interest is that while the proportion of instructional time spent on English or Spanish language arts may vary by program and grade, the total proportion of time allocated to English and Spanish language arts ranges from one-half (50.3%) to roughly two-thirds (67.3%) in all grades within and across programs.

Fragmentation of the educational program is a major concern for all students receiving special support services. One indicator of the quality of services is the extent to which instruction is provided by someone other than the classroom teacher and the extent to which these teachers coordinate their instructional efforts. Over one-half of all teachers by grade within and across programs indicate that their LEP students are taught by another teacher (see Table 83). Slightly more than one-half to over three-fourths of immersion strategy (54.2% to 73.8%) and early-exit (63.0% to 78.8%) teachers indicate that their LEP students are taught the same subject(s) by another teacher either in a pull-out or in-classroom teaching arrangement. Minor grade level differences are noted between these two programs in kindergarten through third grade; however, more early-exit fourth grade teachers (78.4%) than immersion strategy teachers (54.2%) report their students being taught by another teacher. Proportionately more late-exit teachers in grades kindergarten, two and three (88.5%, 79.7% and 80.0%, respectively) than either immersion strategy or early-exit teachers in those grades note that their LEP students are taught by someone else. The majority of the classroom teachers (80.9% to 100.0%), by grade within and across programs, coordinate their instruction with these other teachers who provide instruction to LEP students (see Table 84).

Classroom teachers typically coordinate instruction with these outside teachers through informal discussions and meetings (53.3% to 100.0%) (see Table 85). Teacher coordination also is aided by the curriculum and instructional materials used (47.5% to 100.0%). Project teachers tend to meet with one another on an as-needed basis (18.2% to 100.0%), although teachers also report meeting on a weekly basis (0.0% to 66.7%) (see Table 86). With the exception of second grade, there is a trend among immersion strategy teachers to meet more frequently (i.e., at least once a week) as grade level increases (26.2% to 66.7%). In sum, on average, LEP students in all three programs and all grades are provided instruction by a resource teacher on a regular basis. Coordination of the instructional program between teachers tends to be informal.

What instructional materials do project teachers use and at what difficulty level do they teach?

Teachers in all grades within and across programs provide instructional materials for their LEP students to develop English language arts, Spanish language arts, social studies, math, and science skills. However, there are marked grade level and program differences in the types of materials used.

Except for Spanish language arts, the majority of immersion strategy teachers use materials containing the same content and the same language with their LEP students as are used with English proficient students (see Table 87). The next largest group of immersion strategy teachers report that while they also use English for instruction, the material covered in the content area is expanded to be sensitive to the language needs of LEP students.

Table 82

Percentage of Instructional Time for Each Academic Subject, by Program and Grade

Grade	Content Area	Program Immersion %	 Early-Exit %	 Late-Exit %
K	English Language Arts	61.2	35.4	23.8
	Spanish Language Arts	0.0	28.6	42.7
	Social Studies	9.0	9.8	8.3
	Mathematics	21.8	19.8	17.6
	Science	8.1	6.5	7.6
1	English Language Arts	61.6	38.6	28.3
	Spanish Language Arts	1.2	23.6	39.0
	Social Studies	8.1	7.1	6.6
	Mathematics	22.6	23.5	19.2
	Science	6.5	7.2	7.0
2	English Language Arts	61.3	43.8	27.2
	Spanish Language Arts	0.4	18.9	36.3
	Social Studies	6.7	7.0	9.1
	Mathematics	22.3	22.6	17.9
	Science	9.2	7.7	9.5
3	English Language Arts	58.8	41.6	36.0
	Spanish Language Arts	0.0	17.9	27.8
	Social Studies	7.1	8.6	7.9
	Mathematics	23.9	24.2	20.2
	Science	10.1	7.7	8.1
4	English Language Arts	51.7	50.3	41.7
	Spanish Language Arts	0.0	0.0	21.7
	Social Studies	19.0	15.1	8.7
	Mathematics	16.2	22.2	19.0
	Science	13.1	12.4	8.9

Table 82 (continued)

Grade	Content Area	Immersion %	Program Early-Exit %	Late-Exit %
5	English Language Arts	*	*	36.9
	Spanish Language Arts	*	*	20.9
	Social Studies	*	*	10.7
	Mathematics	*	*	21.0
	Science	*	*	10.5
6	English Language Arts	*	*	37.0
	Spanish Language Arts	*	*	17.6
	Social Studies	*	*	13.9
	Mathematics	*	*	20.4
	Science	*	*	11.2
	Weighted Number of responses	179	180	180

ERIC

Table 83

Percentage of Teachers Reporting that LEP Students are Taught Same Subject(s) by Other Teachers, by Program and Grade

Grade	Program Immersion %	 Early-Exit %	 Late-Exit %
K	73.8	78.8	88.5
1	59.5	65.4	51.6
2	70.9	65.5	79.7
3	57.1	63.0	80.0
4	54.2	78.4	67.6
5	*	*	86.7
6	*	*	80.8
Weighted Number of responses	152	160	163

Table 84

Percentage of Teachers Who Coordinate Instruction with Other Teachers Who Instruct LEPs in Same Subjects, by Program and Grade

Grade	Program Immersion %	 Early-Exit %	 Late-Exit %
K	100.0	100.0	100.0
1	100.0	100.0	100.0
2	100.0	97.7	80.9
3	90.2	100.0	100.0
4	100.0	100.0	85.5
5	*	*	100.0
6	*	*	100.0
Weighted Number of responses	99	109	124

Table 85

Percentage of Teachers Reporting Methods of Coordinating Instruction, by Program and Grade

Grade		Program: Immersion %	Early-Exit %	Late-Exit %
K	Informal discussions/ meetings	100.0	91.6	79.6
	Formal meetings	10.7	56.0	24.7
	Curriculum/materials	47.8	82.0	79.6
	Student Feedback	0.0	21.0	28.7
1	Informal discussions/ meetings	80.8	63.1	81.8
	Formal meetings	37.9	52.0	100.0
	Curriculum/materials	47.5	52.7	100.0
	Student Feedback	10.4	20.6	15.7
2	Informal discussions/ meetings	90.7	83.2	53.3
	Formal meetings	68.0	77.1	71.5
	Curriculum/materials	90.9	78.3	76.9
	Student Feedback	60.6	29.7	47.4
3	Informal discussions/ meetings	76.5	88.5	83.2
	Formal meetings	41.5	59.3	35.5
	Curriculum/materials	70.7	87.8	79.6
	Student Feedback	17.3	54.9	46.7
4	Informal discussions/ meetings	100.0	100.0	68.9
	Formal meetings	100.0	56.4	38.4
	Curriculum/materials	100.0	100.0	100.0
	Student Feedback	66.7	73.4	0.0
5	Informal discussions/ meetings	*	*	64.8
	Formal meetings	*	*	69.3
	Curriculum/materials	*	*	54.5
	Student Feedback	*	*	31.3

Table 85 (continued)

Grade		Immersion %	Program Early-Exit %	Late-Exit %
6	Informal discussions/ meetings	*	*	100.0
	Formal meetings	*	*	53.7
	Curriculum/materials	*	*	77.5
	Student Feedback	*	*	26.9
Range of Weighted Number of responses		48-60	53-61	51-59

Table 86

Percentage of Teachers Reporting Frequency They Meet with Other Project Teachers, by Program and Grade

Grade		Program Immersion %	Early-Exit %	Late-Exit %
K	At least once a week	26.2	51.2	25.0
	At least once a month	28.8	0.0	4.7
	As needed	19.5	48.8	56.9
	Once or twice per year	0.0	0.0	6.7
	Other	25.5	0.0	6.7
1	At least once a week	36.2	35.2	23.1
	At least once a month	12.5	5.9	0.0
	As needed	30.6	50.3	58.4
	Once or twice per year	10.8	0.0	6.3
	Other	9.9	8.6	12.2
2	At least once a week	17.8	13.4	0.0
	At least once a month	11.7	25.6	0.0
	As needed	65.2	43.3	100.0
	Once or twice per year	5.3	8.8	0.0
	Other	0.0	8.9	0.0
3	At least once a week	45.5	41.9	25.3
	At least once a month	0.0	0.0	11.0
	As needed	54.5	58.2	43.1
	Once or twice per year	0.0	0.0	3.6
	Other	0.0	0.0	17.1
4	At least once a week	66.7	56.4	26.0
	At least once a month	0.0	8.4	11.6
	As needed	33.3	18.2	53.8
	Once or twice per year	0.0	16.9	0.0
	Other	0.0	0.0	8.7
5	At least once a week	*	*	50.0
	At least once a month	*	*	16.7
	As needed	*	*	33.3
	Once or twice per year	*	*	0.0
	Other	*	*	0.0
6	At least once a week	*	*	40.7
	At least once a month	*	*	16.5
	As needed	*	*	36.7
	Once or twice per year	*	*	0.0
	Other	*	*	6.2
	Weighted number of responses	81	78	89

Table 87

Percentage of Immersion Strategy Teachers Reporting Types of Materials Used in Class, by Grade

Gr.	Content Area	Same Content Same Lang. %	Same Content Diff Lang. %	Expanded LEP Content Same Lang. %	Diff Content Diff Lang. %
K	Oral English	52.5	0.0	47.6	0.0
	Eng. Reading	60.1	0.0	39.9	0.0
	Eng. Writing	53.2	0.0	46.8	0.0
	Oral Spanish	100.0	0.0	0.0	0.0
	Span. Reading	M	M	M	M
	Span. Writing	M	M	M	M
	Soc. Studies	90.5	0.0	9.5	0.0
	Math	92.0	0.0	8.0	0.0
	Science	90.5	0.0	9.5	0.0
1	Oral English	59.1	2.7	38.3	0.0
	Eng. Reading	59.2	5.2	35.6	0.0
	Eng. Writing	71.5	2.7	25.8	0.0
	Oral Spanish	0.0	0.0	44.1	55.9
	Span. Reading	0.0	0.0	100.0	0.0
	Span. Writing	0.0	0.0	100.0	0.0
	Soc. Studies	82.8	0.0	17.3	0.0
	Math	93.2	0.0	6.8	0.0
	Science	82.8	0.0	17.3	0.0
2	Oral English	69.4	0.0	30.6	0.0
	Eng. Reading	71.2	2.2	26.5	0.0
	Eng. Writing	80.5	2.2	17.3	0.0
	Oral Spanish	42.0	58.0	0.0	0.0
	Span. Reading	42.0	58.0	0.0	0.0
	Span. Writing	56.9	43.1	0.0	0.0
	Soc. Studies	83.0	5.9	11.1	0.0
	Math	89.7	5.9	4.4	0.0
	Science	80.7	6.1	13.2	0.0
3	Oral English	59.3	0.0	35.9	4.8
	Eng. Reading	74.6	0.0	25.4	0.0
	Eng. Writing	79.7	0.0	20.3	0.0
	Oral Spanish	M	M	M	M
	Span. Reading	M	M	M	M
	Span. Writing	M	M	M	M
	Soc. Studies	94.4	0.0	5.6	0.0
	Math	95.2	0.0	4.8	0.0
	Science	94.7	0.0	5.4	0.0

Table 87 (continued)

Gr.	Content Area	Same Content Same Lang. %	Same Content Diff Lang. %	Expanded LEP Content Same Lang. %	Diff Content Diff Lang. %
4	Oral English	50.0	0.0	50.0	0.0
	Eng. Reading	77.1	0.0	22.9	0.0
	Eng. Writing	72.1	0.0	27.9	0.0
	Oral Spanish	M	M	M	M
	Span. Reading	M	M	M	M
	Span. Writing	M	M	M	M
	Soc. Studies	77.1	0.0	22.9	0.0
	Math	72.1	0.0	27.9	0.0
	Science	77.1	0.0	22.9	0.0

Weighted Number of responses for English language arts 133 - 142

Weighted Number of responses for Spanish language arts 5 - 9

Weighted Number of responses for Social Studies, Math and Science 138 - 143

Typically, through the third grade, approximately two-thirds (or more) of early-exit teachers indicate that they use materials with the same content and language as are used with English proficient students when teaching science and English oral, reading, and writing skills (see Table 88). The remaining early-exit teachers in grades kindergarten through three tend to expand the English language arts content to include materials specifically developed for LEP students. In grade four, all of the teachers in the early-exit program report using English language arts materials that are identical to those used with English proficient students. As grade level increases, proportionately more early-exit teachers tend to use the same content and language with LEP students as with English proficient students when teaching math (37.0% to 72.6%) and social studies (37.3% to 100.0%).

Reflecting the early-exit model of primary language instruction, proportionately more early-exit than immersion strategy teachers report providing the same content for LEP students as is used with English proficient students, but they do so in Spanish rather than in English for social studies math and science. Understandably, early-exit teachers also use tailored materials (i.e., different content and language) when teaching Spanish oral, reading, and writing skills to LEP students. Thus, while early-exit teachers tend to use the same materials and language with LEPs as with English proficient students, many such teachers supplement these with special materials to accommodate the needs of the LEP students.

In keeping with their instructional model, late-exit teachers increase their use (albeit unevenly and not to the same extent as early-exit teachers) of materials for English proficient students with their LEP students as grade level increases (see Table 89). Generally across grades, proportionately more late-exit than early-exit or immersion strategy teachers provide their LEP students with Spanish language materials with the same content in science, social studies, and math as the English language materials that are used with English proficient students. In marked contrast to the program models, proportionately more early-exit teachers use different content and Spanish when teaching Spanish oral language, reading, and writing skills than do late-exit teachers.

Teachers in all grades and programs tend to use a combination of commercial and teacher-made instructional materials (see Table 90). While teachers indicate that they tend to use mostly commercial materials supplemented with some teacher-made materials (44.5% to 81.2%), another large proportion of teachers report that their materials are approximately half commercial and half teacher-made (4.5% to 44.8%).

ERIC

Table 88

Percentage of Early-Exit Teachers Reporting Types of Materials Used in Class, by Grade

Gr.	Content Area	Same Content Same Lang. %	Same Content Diff Lang. %	Expanded LEP Content Same Lang. %	Diff Content Diff Lang. %
K	Oral English	30.4	9.8	59.8	0.0
	Eng. Reading	74.6	0.0	0.0	25.4
	Eng. Writing	60.0	19.6	0.0	20.5
	Oral Spanish	0.0	14.0	0.0	86.0
	Span. Reading	0.0	19.6	0.0	80.5
	Span. Writing	0.0	0.0	0.0	100.0
	Soc. Studies	37.3	37.9	24.8	0.0
	Math	37.0	63.0	0.0	0.0
	Science	62.1	37.9	0.0	0.0
1	Oral English	58.6	0.0	41.5	0.0
	Eng. Reading	53.3	4.2	34.5	8.0
	Eng. Writing	71.7	0.0	20.3	8.0
	Oral Spanish	22.0	2.2	4.3	71.5
	Span. Reading	15.7	2.0	3.8	78.6
	Span. Writing	15.7	2.0	3.8	78.6
	Soc. Studies	60.0	27.8	0.0	12.2
	Math	56.9	40.0	0.0	3.1
	Science	64.9	22.5	0.0	12.6
2	Oral English	50.5	5.3	44.2	0.0
	Eng. Reading	67.4	2.[illegible]	29.9	0.0
	Eng. Writing	78.6	7.1	14.2	0.0
	Oral Spanish	7.0	10.5	3.5	79.1
	Span. Reading	9.2	7.7	3.1	80.0
	Span. Writing	3.5	27.7	3.4	65.4
	Soc. Studies	53.1	38.0	3.9	5.1
	Math	58.9	33.8	2.5	4.9
	Science	64.0	24.8	3.8	7.4
3	Oral English	50.2	4.7	31.4	13.7
	Eng. Reading	69.5	0.0	19.5	11.1
	Eng. Writing	61.5	5.3	23.0	10.2
	Oral Spanish	13.6	13.4	0.0	73.0
	Span. Reading	9.3	18.3	0.0	72.4
	Span. Writing	6.4	25.3	0.0	68.3
	Soc. Studies	47.8	41.0	2.1	9.1
	Math	57.9	31.2	2.0	8.9
	Science	60.5	32.7	2.1	4.7

Table 88 (continued)

Gr.	Content Area	Same Content Same Lang. %	Same Content Diff Lang. %	Expanded LEP Content Same Lang. %	Diff Content Diff Lang. %
4	Oral English	100.0	0.0	0.0	0.0
	Eng. Reading	100.0	0.0	0.0	0.0
	Eng. Writing	100.0	0.0	0.0	0.0
	Oral Spanish	0.0	0.0	0.0	100.0
	Span. Reading	0.0	48.1	0.0	51.9
	Span. Writing	0.0	0.0	0.0	100.0
	Soc. Studies	100.0	0.0	0.0	0.0
	Math	72.6	27.5	0.0	0.0
	Science	72.6	27.5	0.0	0.0

Weighted Number of responses 92 - 136

Table 89

Percentage of Late-Exit Teachers Reporting Types of Materials Used in Class, by Grade

Gr.	Content Area	Same Content Same Lang. %	Same Content Diff Lang. %	Expanded LEP Content Same Lang. %	Diff Content Diff Lang. %
K	Oral English	0.0	40.0	60.0	0.0
	Eng. Reading	0.0	50.0	50.0	0.0
	Eng. Writing	0.0	0.0	100.0	0.0
	Oral Spanish	16.7	83.3	0.0	0.0
	Span. Reading	16.7	83.3	0.0	0.0
	Span. Writing	20.0	80.0	0.0	0.0
	Soc. Studies	25.0	75.0	0.0	0.0
	Math	16.7	83.3	0.0	0.0
	Science	16.7	83.3	0.0	0.0
1	Oral English	40.1	4.3	48.4	7.2
	Eng. Reading	36.7	21.5	34.4	7.4
	Eng. Writing	45.8	26.1	11.7	16.4
	Oral Spanish	14.0	28.2	0.0	57.8
	Span. Reading	13.2	26.5	0.0	60.4
	Span. Writing	13.2	26.5	0.0	60.4
	Soc. Studies	20.0	48.6	0.0	31.4
	Math	22.3	68.8	0.0	8.9
	Science	40.8	56.8	0.0	2.5
2	Oral English	45.2	9.8	45.0	0.0
	Eng. Reading	63.5	24.4	12.2	0.0
	Eng. Writing	46.5	17.8	26.8	8.9
	Oral Spanish	25.8	13.9	12.3	48.0
	Span. Reading	26.4	14.4	6.6	52.6
	Span. Writing	31.2	12.9	6.6	49.3
	Soc. Studies	32.2	58.2	0.0	9.6
	Math	25.0	64.5	3.3	7.3
	Science	40.7	51.2	0.0	8.1
3	Oral English	68.6	5.4	20.5	5.4
	Eng. Reading	83.3	5.4	11.3	0.0
	Eng. Writing	87.0	5.4	7.5	0.0
	Oral Spanish	63.9	0.0	16.9	19.3
	Span. Reading	57.1	0.0	16.2	26.8
	Span. Writing	57.1	0.0	16.2	26.8
	Soc. Studies	65.8	28.3	0.0	5.9
	Math	66.8	21.7	5.8	5.8
	Science	75.5	9.5	3.9	11.2

Table 89 (continued)

Gr.	Content Area	Same Content Same Lang. %	Same Content Diff Lang. %	Expanded LEP Content Same Lang. %	Diff Content Diff Lang. %
4	Oral English	45.8	9.4	41.0	3.9
	Eng. Reading.	56.9	9.4	29.9	3.9
	Eng. Writing	68.0	9.4	18.8	3.9
	Oral Spanish	62.2	9.7	0.0	28.1
	Span. Reading	62.2	14.2	0.0	23.7
	Span. Writing	55.7	20.6	0.0	23.7
	Soc. Studies	50.2	38.1	5.9	5.9
	Math	54.2	34.0	5.9	5.9
	Science	56.5	27.6	5.9	10.0
5	Oral English	43.1	9.9	41.3	5.8
	Eng. Reading	54.7	10.0	23.6	11.8
	Eng. Writing	56.5	19.9	23.6	0.0
	Oral Spanish	54.6	10.9	0.0	34.5
	Span. Reading	49.9	12.4	0.0	37.7
	Span. Writing	49.9	21.0	0.0	29.1
	Soc. Studies	50.1	35.1	5.5	9.3
	Math	66.6	22.4	5.5	5.5
	Science	59.4	31.3	0.0	9.3
6	Oral English	50.5	19.1	19.1	11.3
	Eng. Reading	58.3	11.3	22.6	7.8
	Eng. Writing	58.3	11.3	22.6	7.8
	Oral Spanish	33.0	33.0	0.0	34.0
	Span. Reading	33.0	33.0	0.0	34.0
	Span. Writing	36.1	36.1	0.0	27.9
	Soc. Studies	42.6	57.4	0.0	0.0
	Math	42.6	57.4	0.0	0.0
	Science	69.6	22.6	0.0	7.8

Weighted Number of responses 101 - 135

ERIC

Table 90

Percentage of Teachers Reporting Source of Instructional Materials, by Program and Grade

Grade	Origin	Immersion %	Program Early-Exit %	Late-Exit %
K	All commercial	3.7	8.1	0.0
	Most commercial, some teacher-made	49.4	81.2	53.2
	About half commercial, half teacher-made	25.6	7.2	27.8
	Most teacher-made, some commercial	21.3	3.5	19.0
1	All commercial	5.1	8.7	19.2
	Most commercial, some teacher-made	67.1	58.1	50.5
	About half commercial, half teacher-made	21.2	33.2	18.5
	Most teacher-made, some commercial	6.6	0.0	11.8
2	All commercial	11.2	12.0	2.3
	Most commercial, some teacher-made	62.5	64.8	72.2
	About half commercial, half teacher-made	26.3	23.2	4.5
	Most teacher-made, some commercial	0.0	0.0	21.1
3	All commercial	10.7	19.1	12.7
	Most commercial, some teacher-made	44.5	61.0	66.1
	About half commercial, half teacher-made	44.8	10.0	17.8
	Most teacher-made, some commercial	0.0	9.8	3.5
4	All commercial	0.0	34.1	20.3
	Most commercial, some teacher-made	77.1	58.6	54.1
	About half commercial, half teacher-made	22.9	7.4	22.0
	Most teacher-made, some commercial	0.0	0.0	3.6

Table 90 (continued)

Grade	Origin	Program Immersion %	Early-Exit %	Late-Exit %
5	All commercial	*	*	33.2
	Most commercial, some teacher-made	*	*	47.4
	About half commercial, half teacher-made	*	*	10.6
	Most teacher-made, some commercial	*	*	8.8
6	All commercial	*	*	28.5
	Most commercial, some teacher-made	*	*	49.6
	About half commercial, half teacher-made	*	*	21.9
	Most teacher-made, some commercial	*	*	0.0
	Weighted Number of Responses	180	180	177

245

To what extent do LEP and FEP students work at the appropriate grade level? Curiously, proportionately more kindergarten immersion strategy teachers perceive that their LEP students are working at grade level in English skills than teachers in subsequent grades (see Table 91). After a marked drop in these skills from kindergarten to grade one, proportionately more teachers rate their LEP students as having grade level English language skills as grade increases. The majority of immersion strategy teachers indicate that their LEP students are working at grade level in social studies (71.4% to 100.0%), math (80.6% to 100.0%), and science (71.9% to 100.0%).

Early-exit teachers report fluctuations in the skill levels of their LEP students across grades and content areas (see Table 92). While the proportion of teachers who rate LEP students as working at grade level on oral English skills fluctuates somewhat as grade level increases, there is a decline in the proportion who rate them at grade level in English writing and reading skills. The majority of teachers in kindergarten through third grade rate their LEP students as working at grade level in social studies and science, and at grade level in math through fourth grade. At least one-half to all teachers indicate that students' Spanish language skills are at grade level in all grades.

While proportionately more late-exit teachers initially rate their students as performing at grade level in all content areas except English language arts when compared to their counterparts in the other instructional programs, there is also a general decline as grade level increases in the proportion who report their students as having grade level skills (see Table 93).

The apparent decline in the proportion of teachers rating their LEP students as working at grade level may reflect the increased demands on higher skills from students as grade level increases, as well as the characteristics of LEP students remaining in these programs (i.e., slower) and/or LEP students new to the district at higher grades from those LEP students who were re-classified at earlier grades.

Table 91

Percentage of Immersion Strategy Teachers Reporting Level at Which LEP Students are Working, by Grade

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
K	Oral English	40.6	59.4	0.0
	Eng. Reading	53.4	46.7	0.0
	Eng. Writing	19.0	81.0	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	M	M	M
	Span. Writing	M	M	M
	Soc. Studies	24.7	75.4	0.0
	Math	16.0	84.0	0.0
	Science	16.0	84.0	0.0
1	Oral English	69.4	30.6	0.0
	Eng. Reading	74.8	25.2	0.0
	Eng. Writing	74.5	25.5	0.0
	Oral Spanish	33.4	66.7	0.0
	Span. Reading	100.0	0.0	0.0
	Span. Writing	16.3	83.7	0.0
	Soc. Studies	25.4	74.6	0.0
	Math	19.4	80.6	0.0
	Science	25.4	74.6	0.0
2	Oral English	56.6	43.4	0.0
	Eng. Reading	79.2	20.9	0.0
	Eng. Writing	66.6	33.4	0.0
	Oral Spanish	54.4	45.6	0.0
	Span. Reading	72.5	27.5	0.0
	Span. Writing	10.5	89.6	0.0
	Soc. Studies	14.9	82.9	2.3
	Math	6.3	91.4	2.3
	Science	23.4	76.6	0.0

Table 91 (continued)

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
3	Oral English	42.4	52.1	5.5
	Eng. Reading	50.9	49.2	0.0
	Eng. Writing	69.7	30.3	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	M	M	M
	Span. Writing	17.3	82.7	0.0
	Soc. Studies	28.7	71.4	0.0
	Math	8.9	86.3	4.8
	Science	28.1	71.9	0.0
4	Oral English	50.0	50.0	0.0
	Eng. Reading	45.8	54.2	0.0
	Eng. Writing	63.9	36.2	0.0
	Oral Spanish	M	M	M
	Span. Reading	M	M	M
	Span. Writing	M	M	M
	Soc. Studies	0.0	100.0	0.0
	Math	0.0	100.0	0.0
	Science	0.0	100.0	0.0

Weighted Number of responses for English language arts 94 - 141

Weighted Number of responses for Spanish language arts 6 - 50

Weighted Number of responses
for Social Studies, Math and Science 100 - 141

ERIC Full Text Provided by ERIC

Table 92

Percentage of Early-Exit Teachers Reporting Level at Which LEP Students are Working, by Grade

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
K	Oral English	66.4	33.6	0.0
	Eng. Reading	50.4	49.6	0.0
	Eng. Writing	39.7	60.3	0.0
	Oral Spanish	50.2	49.8	0.0
	Span. Reading	36.8	63.2	0.0
	Span. Writing	0.0	100.0	0.0
	Soc. Studies	39.9	60.2	0.0
	Math	29.6	50.6	19.8
	Science	39.9	60.2	0.0
1	Oral English	56.5	43.5	0.0
	Eng. Reading	58.7	38.7	2.6
	Eng. Writing	38.1	57.5	4.4
	Oral Spanish	14.3	80.1	5.6
	Span. Reading	26.8	68.0	5.2
	Span. Writing	23.7	76.3	0.0
	Soc. Studies	5.9	89.5	4.6
	Math	9.2	89.2	1.6
	Science	4.4	93.3	2.3
2	Oral English	55.1	44.9	0.0
	Eng. Reading	55.1	44.9	0.0
	Eng. Writing	38.3	61.7	0.0
	Oral Spanish	33.0	63.9	3.1
	Span. Reading	39.5	60.5	0.0
	Span. Writing	39.5	57.2	3.3
	Soc. Studies	17.1	78.0	5.0
	Math	7.5	92.5	0.0
	Science	27.2	72.8	0.0
3	Oral English	66.3	33.7	0.0
	Eng. Reading	74.2	25.8	0.0
	Eng. Writing	59.5	40.6	0.0
	Oral Spanish	7.0	93.0	0.0
	Span. Reading	32.6	67.4	0.0
	Span. Writing	41.2	58.8	0.0
	Soc. Studies	9.1	90.9	0.0
	Math	13.1	87.0	0.0
	Science	12.3	87.7	0.0

Table 92 (continued)

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
4	Oral English	54.2	45.8	0.0
	Eng. Reading	84.9	15.1	0.0
	Eng. Writing	67.2	32.8	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	0.0	100.0	0.0
	Span. Writing	50.0	50.0	0.0
	Soc. Studies	69.5	30.5	0.0
	Math	16.0	84.1	0.0
	Science	69.5	30.5	0.0

Weighted Number of responses 94 - 139

ERIC

Table 93

Percentage of Late-Exit Teachers Reporting Level at Which LEP Students are Working, by Grade

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
K	Oral English	16.7	83.3	0.0
	Eng. Reading	50.0	50.0	0.0
	Eng. Writing	66.7	33.3	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	0.0	100.0	0.0
	Span. Writing	0.0	100.0	0.0
	Soc. Studies	0.0	100.0	0.0
	Math	0.0	100.0	0.0
	Science	0.0	100.0	0.0
1	Oral English	69.9	30.1	0.0
	Eng. Reading	100.0	0.0	0.0
	Eng. Writing	67.8	24.2	8.1
	Oral Spanish	13.2	79.4	7.4
	Span. Reading	18.5	81.5	0.0
	Span. Writing	14.1	85.9	0.0
	Soc. Studies	4.4	95.7	0.0
	Math	7.2	92.8	0.0
	Science	5.9	94.1	0.0
2	Oral English	63.6	36.4	0.0
	Eng. Reading	92.1	7.9	0.0
	Eng. Writing	39.2	60.8	0.0
	Oral Spanish	17.4	82.6	0.0
	Span. Reading	33.7	66.3	0.0
	Span. Writing	32.3	67.8	0.0
	Soc. Studies	9.3	87.4	3.2
	Math	17.3	82.7	0.0
	Science	16.6	83.4	0.0
3	Oral English	52.0	48.0	0.0
	Eng. Reading	69.4	30.6	0.0
	Eng. Writing	70.5	29.5	0.0
	Oral Spanish	11.0	76.1	12.9
	Span. Reading	27.1	66.7	6.2
	Span. Writing	43.8	50.0	6.2
	Soc. Studies	5.6	94.4	0.0
	Math	11.9	88.1	0.0
	Science	10.9	89.1	0.0

Table 93 (continued)

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
4	Oral English	94.7	5.3	0.0
	Eng. Reading	87.3	12.8	0.0
	Eng. Writing	51.6	48.4	0.0
	Oral Spanish	29.2	70.8	0.0
	Span. Reading	34.9	65.1	0.0
	Span. Writing	41.3	58.8	0.0
	Soc. Studies	16.4	83.6	0.0
	Math	11.0	89.0	0.0
	Science	24.9	75.1	0.0
5	Oral English	66.8	33.2	0.0
	Eng. Reading	74.4	25.6	0.0
	Eng. Writing	53.6	39.8	6.6
	Oral Spanish	29.0	71.0	0.0
	Span. Reading	47.0	53.1	0.0
	Span. Writing	39.0	61.0	0.0
	Soc. Studies	31.4	68.6	0.0
	Math	32.7	67.3	0.0
	Science	56.0	44.0	0.0
6	Oral English	73.0	27.0	0.0
	Eng. Reading	73.0	15.7	11.3
	Eng. Writing	92.2	7.8	0.0
	Oral Spanish	8.5	79.3	12.3
	Span. Reading	27.0	61.8	11.3
	Span. Writing	61.8	27.0	11.3
	Soc. Studies	27.0	73.0	0.0
	Math	7.8	92.2	0.0
	Science	27.0	73.0	0.0

Weighted Number of responses 100 - 135

In general, the majority of immersion strategy teachers report that their FEP students are working at grade level in English language arts, social studies, math and science (see Table 94). Consistent with the program model, very few teachers rate these students on Spanish language arts. With one exception, the majority of early-exit teachers indicate that their FEP students are working at grade level in all content areas (see Table 95). Proportionately more early-exit teachers report them as working below grade level in English language arts than

ERIC

in social studies, math and science. More than half of late-exit teachers report that their FEP students are working at grade level or above it (see Table 96). More teachers in the late-exit program indicate that these students work above grade level in Spanish language arts than in other content areas.

How do project teachers rate their programs?

Teachers were asked to rate various aspects of their program on a scale of 1 (very strong) to 4 (most in need of improvement). On the average, immersion strategy teachers rate teaching staff, instructional materials, and curriculum as the strongest aspects of their program (see Table 97). While early-exit teachers also note that the teaching staff is one of the strongest qualities of their program, they also rate site leadership as a strong feature of their program. Late-exit teachers also tend to rate teaching staff and curriculum as project strengths.

Teachers consider certain services as absolutely necessary for their program to function (see Table 98). Teachers in all three programs and grades rate training, special curriculum materials, meetings with other teachers and parent community participation as important for program functioning. In addition to this list, proportionately more early-exit and late-exit teachers note the importance of bilingual aides than do immersion strategy teachers. In all three programs, after kindergarten, higher percentages of teachers report that release time is a necessity to the success of their respective programs.

Table 94

Percentage of Immersion Strategy Teachers Reporting Level at Which FEP Students are Working, by Grade

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
K	Oral English	12.9	74.3	12.9
	Eng. Reading	28.2	71.8	0.0
	Eng. Writing	14.8	67.6	17.6
	Oral Spanish	M	M	M
	Span. Reading	M	M	M
	Span. Writing	M	M	M
	Soc. Studies	0.0	71.8	28.2
	Math	0.0	84.6	15.4
	Science	0.0	84.6	15.4
1	Oral English	33.1	63.0	3.9
	Eng. Reading	30.7	65.4	3.9
	Eng. Writing	51.4	48.6	0.0
	Oral Spanish	100.0	0.0	0.0
	Span. Reading	100.0	0.0	0.0
	Span. Writing	10.6	89.4	0.0
	Soc. Studies	4.9	95.1	0.0
	Math	9.8	85.4	4.9
	Science	7.8	95.2	0.0
2	Oral English	23.9	76.1	0.0
	Eng. Reading	44.3	55.7	0.0
	Eng. Writing	63.8	36.2	0.0
	Oral Spanish	M	M	M
	Span. Reading	0.0	100.0	0.0
	Span. Writing	13.1	86.9	0.0
	Soc. Studies	18.0	82.0	0.0
	Math	14.3	85.7	0.0
	Science	17.5	82.5	0.0
3	Oral English	16.4	83.6	0.0
	Eng. Reading	19.0	81.0	0.0
	Eng. Writing	29.4	70.6	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	M	M	M
	Span. Writing	0.0	100.0	0.0
	Soc. Studies	14.9	85.1	0.0
	Math	0.0	100.0	0.0
	Science	9.2	90.8	0.0

ERIC

Table 94 (continued)

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
4	Oral English	0.0	100.0	0.0
	Eng. Reading	27.9	72.1	0.0
	Eng. Writing	27.9	72.1	0.0
	Oral Spanish	M	M	M
	Span. Reading	M	M	M
	Span. Writing	M	M	M
	Soc. Studies	0.0	100.0	0.0
	Math	0.0	100.0	C.0
	Science	0.0	100.0	0.0

Weighted Number of responses for English language arts 57 - 83

Weighted Number of responses for Spanish language arts 2 - 26

Weighted Number of responses for Social Studies, Math and Science 62 - 82

Table 95

Percentage of Early-Exit Teachers Reporting Level at Which FEP Students are Working, by Grade

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
K	Oral English	17.1	83.0	0.0
	Eng. Reading	33.3	66.7	0.0
	Eng. Writing	50.0	50.0	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	M	M	M
	Span. Writing	M	M	M
	Soc. Studies	17.1	83.0	0.0
	Math	0.0	100.0	0.0
	Science	17.1	83.0	0.0
1	Oral English	17.3	79.7	3.0
	Eng. Reading	27.5	69.7	2.8
	Eng. Writing	27.1	72.9	0.0
	Oral Spanish	0.0	89.8	10.2
	Span. Reading	0.0	85.6	14.4
	Span. Writing	7.8	92.2	0.0
	Soc. Studies	5.5	91.7	2.8
	Math	0.0	91.2	8.8
	Science	0.0	93.6	6.4
2	Oral English	22.7	73.4	4.0
	Eng. Reading	21.6	68.7	9.7
	Eng. Writing	34.6	60.2	5.2
	Oral Spanish	26.4	73.6	0.0
	Span. Reading	44.9	55.1	0.0
	Span. Writing	19.5	80.5	0.0
	Soc. Studies	2.0	94.1	3.9
	Math	2.0	94.1	3.9
	Science	0.0	95.5	4.5
3	Oral English	26.6	73.4	0.0
	Eng. Reading	22.9	71.4	5.7
	Eng. Writing	39.8	60.2	0.0
	Oral Spanish	33.5	66.5	0.0
	Span. Reading	79.4	20.6	0.0
	Span. Writing	11.2	88.9	0.0
	Soc. Studies	8.7	91.4	0.0
	Math	2.9	97.1	0.0
	Science	16.7	83.4	0.0

Table 95 (continued)

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
4	Oral English	26.2	73.8	0.0
	Eng. Reading	30.0	70.0	0.0
	Eng. Writing	41.7	58.4	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	0.0	100.0	0.0
	Span. Writing	0.0	100.0	0.0
	Soc. Studies	23.7	76.3	0.0
	Math	0.0	100.0	0.0
	Science	23.7	76.3	0.0

Weighted number of responses 21 - 90

ERIC

Table 96

Percentage of Late-Exit Teachers Reporting Level at Which FEP Students are Working, by Grade

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
K	Oral English	25.0	75.0	0.0
	Eng. Reading	0.0	100.0	0.0
	Eng. Writing	0.0	100.0	0.0
	Oral Spanish	0.0	100.0	0.0
	Span. Reading	0.0	100.0	0.0
	Span. Writing	0.0	100.0	0.0
	Soc. Studies	0.0	100.0	0.0
	Math	0.0	100.0	0.0
	Science	0.0	100.0	0.0
1	Oral English	4.1	95.9	0.0
	Eng. Reading	46.0	48.2	5.8
	Eng. Writing	32.3	60.5	7.2
	Oral Spanish	8.2	83.6	8.2
	Span. Reading	0.0	91.8	8.2
	Span. Writing	8.2	91.8	0.0
	Soc. Studies	0.0	100.0	0.0
	Math	0.0	93.3	6.7
	Science	0.0	100.0	0.0
2	Oral English	27.2	67.2	5.6
	Eng. Reading	58.5	41.5	0.0
	Eng. Writing	23.9	70.5	5.6
	Oral Spanish	25.1	74.9	0.0
	Span. Reading	23.9	66.7	9.5
	Span. Writing	18.6	81.4	0.0
	Soc. Studies	8.5	91.5	0.0
	Math	9.3	90.7	0.0
	Science	0.0	100.0	0.0
3	Oral English	10.3	83.6	6.1
	Eng. Reading	42.2	51.7	6.1
	Eng. Writing	49.4	44.5	6.1
	Oral Spanish	15.5	70.5	14.0
	Span. Reading	18.1	60.2	21.8
	Span. Writing	28.7	49.5	21.8
	Soc. Studies	0.0	94.1	5.9
	Math	5.8	82.6	11.7
	Science	0.0	94.2	5.8

Table 96 (continued)

Grade	Content Area	Below Grade Level %	At Grade Level %	Above Grade Level %
4	Oral English	37.3	56.5	6.1
	Eng. Reading	38.2	55.9	6.0
	Eng. Writing	19.1	74.7	6.2
	Oral Spanish	4.7	81.9	13.4
	Span. Reading	18.1	61.8	20.1
	Span. Writing	29.4	57.2	13.4
	Soc. Studies	9.8	84.5	5.8
	Math	12.3	81.6	6.1
	Science	5.6	94.4	0.0
5	Oral English	25.0	68.8	6.2
	Eng. Reading	34.2	65.8	0.0
	Eng. Writing	31.9	64.6	3.5
	Oral Spanish	34.4	59.0	6.6
	Span. Reading	47.5	52.5	0.0
	Span. Writing	23.3	76.7	0.0
	Soc. Studies	24.1	69.9	6.0
	Math	17.5	82.5	0.0
	Science	37.3	62.7	0.0
6	Oral English	15.7	73.0	11.3
	Eng. Reading	12.2	70.6	17.1
	Eng. Writing	31.6	68.4	0.0
	Oral Spanish	16.3	60.2	23.5
	Span. Reading	39.9	37.9	22.2
	Span. Writing	53.5	34.0	12.5
	Soc. Studies	6.4	79.8	13.9
	Math	9.6	80.8	9.6
	Science	9.2	81.7	9.2

Weighted number of responses 84 - 112

ERIC

Table 97

Mean Teacher Ratings of Program Strengths and Weaknesses+, by Program and Grade

Grade		Immersion	Program Early-Exit	Late-Exit
K	Materials	1.9	2.3	2.0
	Diagnostic Procedures	2.2	2.4	2.7
	Staff Training	2.0	2.2	3.0
	Curriculum	2.3	2.0	2.0
	Project Leadership	2.4	1.9	2.7
	Site Leadership	2.0	2.1	2.6
	Teaching Staff	1.7	2.4	1.8
	Parent Involvement	2.3	3.0	1.8
1	Materials	1.8	2.0	2.4
	Diagnostic Procedures	2.0	2.2	2.5
	Staff Training	1.9	2.1	2.4
	Curriculum	1.9	2.0	2.6
	Project Leadership	1.9	2.1	2.4
	Site Leadership	1.8	2.0	2.4
	Teaching Staff	1.6	1.9	2.1
	Parent Involvement	2.6	2.8	3.2
2	Materials	1.9	2.0	2.2
	Diagnostic Procedures	2.1	1.9	2.4
	Staff Training	2.0	2.0	2.3
	Curriculum	1.9	2.0	2.1
	Project Leadership	1.9	2.0	2.4
	Site Leadership	2.0	1.9	2.4
	Teaching Staff	1.6	1.9	2.2
	Parent Involvement	2.8	2.8	2.9
3	Materials	2.3	2.1	2.3
	Diagnostic Procedures	2.1	2.3	2.1
	Staff Training	2.2	1.9	2.2
	Curriculum	2.0	2.0	2.0
	Project Leadership	2.1	1.9	2.3
	Site Leadership	1.8	1.9	2.3
	Teaching Staff	1.6	1.7	2.1
	Parent Involvement	2.7	3.0	2.7

Grade		Immersion	Program Early-Exit	Late-Exit
4	Materials	1.8	2.2	2.5
	Diagnostic Procedures	2.5	2.3	2.1
	Staff Training	2.1	1.8	2.3
	Curriculum	1.8	2.0	2.2
	Project Leadership	2.0	2.0	2.3
	Site Leadership	2.0	1.8	2.3
	Teaching Staff	1.4	1.6	2.0
	Parent Involvement	3.2	3.3	2.9
5	Materials	*	*	2.5
	Diagnostic Procedures	*	*	2.6
	Staff Training	*	*	2.5
	Curriculum	*	*	2.4
	Project Leadership	*	*	2.6
	Site Leadership	*	*	2.6
	Teaching Staff	*	*	2.4
	Parent Involvement	*	*	3.0
6	Materials	*	*	2.3
	Diagnostic Procedures	*	*	2.3
	Staff Training	*	*	1.8
	Curriculum	*	*	2.2
	Project Leadership	*	*	2.2
	Site Leadership	*	*	1.9
	Teaching Staff	*	*	1.9
	Parent Involvement	*	*	2.8
	Weighted Number of Responses	138 - 144	140 - 147	139 - 147

+ Teachers rated program strengths and weaknesses on a 4-point scale ranging from 1 (very strong) to 4 (most in need of improvement).

Table 98

Percentage of Teachers Rating Absolutely Necessary, by Program and Grade

Gr.		Immersion %	Program Early-Exit %	Late-Exit %
K	Training	75.5	74.5	81.0
	Special curriculum materials	70.5	84.3	85.9
	Supervision	42.0	37.3	68.1
	Release time	46.9	54.5	34.4
	Homogeneous grouping of students by language proficiency	22.6	55.3	72.2
	Aides (monolingual English)	11.2	20.1	0.0
	Bilingual aides	63.6	90.1	94.1
	Meetings with other teachers	85.2	73.5	77.9
	Parent/community participation	81.9	75.7	76.9
1	Training	89.8	84.6	93.7
	Special curriculum materials	89.2	93.3	95.7
	Supervision	61.8	47.8	42.6
	Release time	66.7	81.6	67.5
	Homogeneous grouping of students by language proficiency	52.9	56.5	91.8
	Aides (monolingual English)	16.9	26.5	9.0
	Bilingual aides	82.1	90.3	86.8
	Meetings with other teachers	85.7	65.4	87.1
	Parent/community participation	73.8	71.6	82.2
2	Training	96.5	83.9	74.6
	Special curriculum materials	84.8	82.9	91.1
	Supervision	70.7	54.8	64.3
	Release time	84.4	79.6	92.5
	Homogeneous grouping of students by language proficiency	49.2	63.2	95.1
	Aides (monolingual English)	26.5	20.6	3.2
	Bilingual aides	77.0	77.5	97.7
	Meetings with other teachers	79.1	82.0	76.6
	Parent/community participation	84.8	85.8	96.9

Table 98 (continued)

Gr.		Immersion %	Program Early-Exit %	Late-Exit %
3	Training	100.0	92.1	78.0
	Special curriculum materials	85.8	84.4	83.1
	Supervision	63.1	74.8	68.1
	Release time	94.9	80.7	66.7
	Homogeneous grouping of students by language proficiency	50.0	58.5	77.6
	Aides (monolingual English)	33.5	31.6	20.3
	Bilingual aides	72.7	84.4	77.3
	Meetings with other teachers	78.6	74.5	74.9
	Parent/community participation	80.7	88.2	89.9
4	Training	100.0	100.0	80.1
	Special curriculum materials	77.1	92.6	84.8
	Supervision	77.1	79.2	61.0
	Release time	100.0	85.8	66.2
	Homogeneous grouping of students by language proficiency	18.1	21.6	73.4
	Aides (monolingual English)	0.0	14.8	42.2
	Bilingual aides	45.8	56.5	71.9
	Meetings with other teachers	81.9	79.2	85.2
	Parent/community participation	81.9	86.0	86.4
5	Training	*	*	87.6
	Special curriculum materials	*	*	89.3
	Supervision	*	*	75.4
	Release time	*	*	71.1
	Homogeneous grouping of students by language proficiency	*	*	75.6
	Aides (monolingual English)	*	*	21.0
	Bilingual aides	*	*	78.1
	Meetings with other teachers	*	*	92.1
	Parent/community participation	*	*	97.3
6	Training	*	*	70.7
	Special curriculum materials	*	*	93.7
	Supervision	*	*	54.4
	Release time	*	*	75.3
	Homogeneous grouping of students by language proficiency	*	*	59.8
	Aides (monolingual English)	*	*	19.0
	Bilingual aides	*	*	64.3
	Meetings with other teachers	*	*	73.4
	Parent/community participation	*	*	84.5
	Weighted Number of Responses	123-180	120-180	120-179

Do project teachers understand the goals and procedures of their programs?

The existence and communication of skills continuua among and by teachers is important to their understanding of the goals and procedures of their respective instructional programs. Overall, across programs and grades, the majority of teachers report using existing written skills continuua in all English content areas; there are few exceptions. However, each instructional program has teachers who report that such continuua in English content do not exist, or who note that while continuua do exist, they do not use them (see Tables 99, 100, and 101). The proportion of such teachers varies by grade level and subject area within program. This suggests that while such criteria do in fact exist within each program, they are not communicated effectively to teachers, or that there are teachers who apparently have goals and procedures that may be inconsistent with those established by their respective school districts. These data may point to a need within each program for greater articulation and coordination with respect to program goals and procedures for each of the content areas when taught in English.

Not surprisingly, the majority of immersion strategy teachers across grades report that there are no continuua for teaching the various content areas in Spanish, or, if they do exist, that they do not use them (see Table 102). While the majority of early-exit and late-exit teachers report that there are continuua for teaching the content areas in Spanish, many also indicate either that they do not use them or that they do not exist (see Tables 103 and 104). This is reported most frequently in science and social studies. As before, these data may suggest a need for greater articulation and coordination of instructional goals and procedures within the early- and late-exit programs.

Table 99

Percentage of Immersion Strategy Teachers Reporting Existence/Use of Written Skills Continua in English, by Grade

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
K	Eng. Reading	100.0	0.0	0.0
	Oral English	84.0	0.0	16.0
	English Writing	100.0	0.0	0.0
	Math in English	66.7	16.0	17.4
	Science in English	66.7	0.0	33.3
	Soc. Studies in English	66.7	0.0	33.3
1	Eng. Reading	92.9	5.1	2.0
	Oral English	83.1	14.4	2.6
	English Writing	90.3	5.1	4.6
	Math in English	92.9	5.1	2.0
	Science in English	88.2	2.6	9.3
	Soc. Studies in English	88.2	5.1	6.7
2	Eng. Reading	86.7	13.3	0.0
	Oral English	75.4	15.5	9.1
	English Writing	100.0	0.0	0.0
	Math in English	91.1	6.7	2.2
	Science in English	86.4	11.4	2.3
	Soc. Studies in English	86.4	11.4	2.3
3	Eng. Reading	84.8	5.1	10.2
	Oral English	62.8	25.9	11.3
	English Writing	95.2	0.0	4.8
	Math in English	84.8	5.1	10.2
	Science in English	74.6	10.2	15.3
	Soc. Studies in English	73.5	10.6	15.9

ERIC

Table 99 (continued)

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
4	Eng. Reading	100.0	0.0	0.0
	Oral English	0.0	54.2	45.8
	English Writing	100.0	0.0	0.0
	Math in English	81.9	18.1	0.0
	Science in English	77.1	0.0	22.9
	Soc. Studies in English	77.1	0.0	22.9

Weighted number of responses 140 - 144

Table 100

Percentage of Early-Exit Teachers Reporting Existence/Use of Written Skills Continuua in English, by Grade

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
K	Eng. Reading	100.0	0.0	0.0
	Oral English	90.2	0.0	9.8
	English Writing	100.0	0.0	0.0
	Math in English	86.0	0.0	14.0
	Science in English	60.5	19.8	19.7
	Soc. Studies in English	60.5	19.8	19.7
1	Eng. Reading	93.6	3.3	3.2
	Oral English	88.2	0.0	11.8
	English Writing	100.0	0.0	0.0
	Math in English	87.0	6.6	6.4
	Science in English	84.7	3.2	12.2
	Soc. Studies in English	90.9	0.0	9.1
2	Eng. Reading	86.0	11.2	2.9
	Oral English	83.3	11.2	5.6
	English Writing	89.6	7.8	2.6
	Math in English	80.4	8.5	11.1
	Science in English	73.4	10.6	15.9
	Soc. Studies in English	70.0	10.9	19.1
3	Eng. Reading	96.0	4.1	0.0
	Oral English	78.4	11.9	9.7
	English Writing	96.0	4.1	0.0
	Math in English	86.2	8.0	5.9
	Science in English	91.5	8.5	0.0
	Soc. Studies in English	91.5	8.5	0.0

Table 100 (continued)

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
4	Eng. Reading	100.0	0.0	0.0
	Oral English	76.1	23.9	0.0
	English Writing	100.0	0.0	0.0
	Math in English	100.0	0.0	0.0
	Science in English	100.0	0.0	0.0
	Soc. Studies in English	86.8	13.2	0.0

Weighted number of responses 130 - 140

Table 101

Percentage of Late-Exit Teachers Reporting Existence/Use of Written Skills Continuum in English, by Grade

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
K	Eng. Reading	16.7	50.0	33.3
	Oral English	83.3	16.7	0.0
	English Writing	33.3	50.0	16.7
	Math in English	0.0	50.0	50.0
	Science in English	16.7	50.0	33.3
	Soc. Studies in English	16.7	50.0	33.3
1	Eng. Reading	33.9	28.7	37.4
	Oral English	61.7	7.4	30.9
	English Writing	60.5	19.7	19.7
	Math in English	47.1	19.0	33.9
	Science in English	40.7	29.5	29.7
	Soc. Studies in English	54.7	15.6	29.7
2	Eng. Reading	78.3	8.9	12.8
	Oral English	88.3	4.8	6.9
	English Writing	77.5	7.5	15.0
	Math in English	62.9	28.7	8.5
	Science in English	56.9	20.4	22.7
	Soc. Studies in English	53.3	22.1	24.6
3	Eng. Reading	79.9	14.6	5.4
	Oral English	77.1	13.7	9.2
	English Writing	94.6	5.4	0.0
	Math in English	79.9	14.6	5.4
	Science in English	78.3	16.3	5.4
	Soc. Studies in English	72.0	22.4	5.6

Table 101 (continued)

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
4	Eng. Reading	79.6	15.3	5.1
	Oral English	60.5	15.5	24.0
	English Writing	81.5	8.0	10.5
	Math in English	82.0	18.0	0.0
	Science in English	61.0	5.7	33.2
	Soc. Studies in English	59.4	14.3	26.3
5	Eng. Reading	96.2	3.8	0.0
	Oral English	75.2	13.5	11.3
	English Writing	77.2	17.6	5.2
	Math in English	81.4	13.1	5.5
	Science in English	56.0	21.1	22.9
	Soc. Studies in English	63.2	25.8	11.0
6	Eng. Reading	79.8	0.0	20.2
	Oral English	39.2	27.0	33.9
	English Writing	82.6	7.1	10.3
	Math in English	72.8	7.0	20.2
	Science in English	55.6	33.5	10.9
	Soc. Studies in English	82.9	7.0	10.1

Weighted number of responses 119 - 130

Table 102

Percentage of Immersion Strategy Teachers Reporting Existence/Use of Written Skills Continua in Spanish, by Grade

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
K	Span. Reading	0.0	29.3	70.7
	Oral Spanish	0.0	29.3	70.7
	Spanish Writing	0.0	19.0	81.0
	Math in Spanish	0.0	19.0	81.0
	Science in Spanish	0.0	19.0	81.0
	Soc. Studies in Spanish	0.0	19.0	81.0
1	Span. Reading	0.0	30.9	69.1
	Oral Spanish	0.0	26.7	73.3
	Spanish Writing	0.0	26.7	73.3
	Math in Spanish	0.0	26.7	73.3
	Science in Spanish	0.0	24.2	75.8
	Soc. Studies in Spanish	0.0	25.1	74.9
2	Span. Reading	10.2	27.8	62.0
	Oral Spanish	10.2	27.8	62.0
	Spanish Writing	10.2	27.8	62.0
	Math in Spanish	10.2	27.8	62.0
	Science in Spanish	17.5	21.7	60.7
	Soc. Studies in Spanish	10.2	21.7	68.1
3	Span. Reading	0.0	59.4	40.6
	Oral Spanish	0.0	51.4	48.6
	Spanish Writing	0.0	50.0	50.0
	Math in Spanish	0.0	50.0	50.0
	Science in Spanish	18.8	27.7	53.6
	Soc. Studies in Spanish	0.0	27.7	72.3

ERIC

Table 102 (continued)

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
4	Span. Reading	0.0	100.0	0.0
	Oral Spanish	0.0	100.0	0.0
	Spanish Writing	0.0	100.0	0.0
	Math in Spanish	0.0	100.0	0.0
	Science in Spanish	0.0	100.0	0.0
	Soc. Studies in Spanish	0.0	100.0	0.0

Weighted number of responses 68 - 71

Table 103

Percentage of Early-Exit Teachers Reporting Existence/Use of Written Skills Continuua in Spanish, by Grade

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
K	Span. Reading	80.2	19.8	0.0
	Oral Spanish	50.5	19.8	29.7
	Spanish Writing	80.2	19.8	0.0
	Math in Spanish	56.0	22.0	22.0
	Science in Spanish	10.3	30.1	59.6
	Soc. Studies in Spanish	10.3	30.1	59.6
1	Span. Reading	88.9	9.5	1.6
	Oral Spanish	60.3	18.1	21.6
	Spanish Writing	65.7	16.3	17.9
	Math in Spanish	73.1	15.8	11.1
	Science in Spanish	21.1	19.4	59.5
	Soc. Studies in Spanish	22.7	22.9	54.4
2	Span. Reading	79.2	8.4	12.5
	Oral Spanish	64.1	17.8	18.1
	Spanish Writing	64.3	15.6	20.2
	Math in Spanish	74.0	11.6	14.5
	Science in Spanish	38.3	22.0	39.7
	Soc. Studies in Spanish	38.2	24.1	37.7
3	Span. Reading	81.5	15.6	2.9
	Oral Spanish	75.2	15.7	9.1
	Spanish Writing	61.8	22.5	15.7
	Math in Spanish	69.1	19.0	12.0
	Science in Spanish	47.3	20.6	32.0
	Soc. Studies in Spanish	53.1	20.6	26.3

Table 103 (continued)

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
4	Span. Reading	0.0	69.1	30.9
	Oral Spanish	40.0	41.5	18.6
	Spanish Writing	0.0	69.1	30.9
	Math in Spanish	0.0	41.5	58.5
	Science in Spanish	40.0	41.5	18.6
	Soc. Studies in Spanish	0.0	41.5	58.5

Weighted number of responses 109 - 116

Table 104

Percentage of Late-Exit Teachers Reporting Existence/Use of Written Skills Continuua in Spanish, by Grade

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
K	Span. Reading	83.3	0.0	16.7
	Oral Spanish	100.0	0.0	0.0
	Spanish Writing	100.0	0.0	0.0
	Math in Spanish	66.7	0.0	33.3
	Science in Spanish	83.3	0.0	16.7
	Soc. Studies in Spanish	83.3	0.0	16.7
1	Span. Reading	72.0	3.0	25.0
	Oral Spanish	67.4	0.0	32.6
	Spanish Writing	62.3	3.0	34.7
	Math in Spanish	57.9	7.4	34.7
	Science in Spanish	57.9	3.1	39.0
	Soc. Studies in Spanish	51.5	0.0	48.5
2	Span. Reading	94.7	0.0	5.3
	Oral Spanish	84.4	0.0	15.6
	Spanish Writing	82.2	2.6	15.2
	Math in Spanish	81.3	3.5	15.2
	Science in Spanish	84.0	0.0	16.0
	Soc. Studies in Spanish	83.0	3.5	13.5
3	Span. Reading	79.2	3.9	16.9
	Oral Spanish	66.5	4.1	29.4
	Spanish Writing	58.3	3.9	37.8
	Math in Spanish	50.5	3.9	45.6
	Science in Spanish	56.1	4.1	39.8
	Soc. Studies in Spanish	62.5	4.1	33.5

Table 104 (continued)

Gr.	Content Area	Exists & Use It %	Exists but Don't Use it %	Doesn't Exist %
4	Span. Reading	69.0	10.0	21.0
	Oral Spanish	48.0	11.6	40.4
	Spanish Writing	54.9	2.7	42.4
	Math in Spanish	60.1	6.4	33.5
	Science in Spanish	37.9	13.4	48.8
	Soc. Studies in Spanish	33.0	13.4	53.6
5	Span. Reading	76.0	0.0	24.1
	Oral Spanish	56.2	7.0	36.9
	Spanish Writing	52.8	16.5	30.7
	Math in Spanish	68.1	4.3	27.7
	Science in Spanish	67.7	0.0	32.3
	Soc. Studies in Spanish	67.2	5.6	27.2
6	Span. Reading	58.0	24.6	17.4
	Oral Spanish	47.7	21.4	30.9
	Spanish Writing	48.4	8.2	43.5
	Math in Spanish	51.4	7.7	41.0
	Science in Spanish	46.7	7.7	45.6
	Soc. Studies in Spanish	47.3	21.4	31.3

Weighted number of responses 133 - 135

Teachers were asked about their views of Spanish use in the classroom under various situations. Immersion strategy teachers across grades feel that Spanish should be used for instruction or social interaction with students only occasionally, to clarify, or as a last resort (see Tables 105 and 106, respectively). Similarly, while the majority of immersion strategy teachers feel that they should accept the use of Spanish among LEP students during instruction and for social interaction, few would encourage this, and some would discourage its use (see Tables 107 and 108). While early-exit kindergarten and first grade teachers indicate that Spanish should be used regularly for

ERIC

formal instruction or social interaction, or regularly for support and explanation, teachers in later grades feel that Spanish should be used only occasionally, to clarify, or as a last resort. While more early-exit teachers would accept the use of Spanish among LEP students during instruction and for social interaction, some would discourage its use. In contrast to the other two programs, the majority of late-exit teachers feel that Spanish should be used regularly for formal instruction and social interaction. Proportionately more late-exit teachers than either immersion strategy or early-exit teachers indicate that they encourage the use of Spanish among LEP students during instruction and for social interaction. In sum, late-exit teachers consistently report a greater role for the use of Spanish in instruction and social interaction than either immersion strategy or early-exit teachers. These views are consistent with the instructional model for each program.

Teachers also were asked for the goals of their programs. There is almost unanimous agreement at each grade level within each program that the goals for each program are to develop English, master the various content areas, and social integration (see Table 109). A large proportion of teachers also support the maintenance of Hispanic culture. The two areas of greatest disagreement are that of development of Spanish language skills and success in an English-only environment. These differences reflect the instructional model of each program. Almost no immersion strategy teacher feels that development of Spanish language skills is a program goal. However, three-fourths or more feel that success in English-only programs is a goal. The proportion of early-exit teachers reporting that development of Spanish language skills is a goal declines while the proportion of teachers rating success in an English-only program as a goal remains more stable as grade level increases. The majority of late-exit teachers across grade levels hold that development of Spanish language skills is an important goal. The proportion of late-exit teachers reporting that success in an English-only program is a goal increases from one-third in kindergarten to two-thirds in grade six, reflecting increased

ERIC

teacher concern at the time when the student must be mainstreamed. Once again, these views are consistent with the instructional model of each program.

Teachers were asked to rank their agreement or disagreement with certain statements according to whether they strongly agree, somewhat agree, somewhat disagree, or strongly disagree. To facilitate comparisons between programs, Table 110 shows the percentage of teachers agreeing strongly or somewhat with each statement. Large proportions of late-exit teachers, followed by those in early-exit programs, feel strongly that LEP students need to develop skills in their primary language similar to the skills they develop in English, that they learn to read English best if they are first taught to read in their primary language, and how well LEP students know their primary language is important in deciding how or what to teach them in school. Markedly fewer immersion strategy teachers hold these views. Immersion strategy teachers, on the other hand, contend that LEP students who have learned good conversational English are ready for content area instruction given entirely in English, that a teacher best uses a LEP student's primary language to support primary instruction given in English, rather than using it as a primary language of instruction, and that the main purpose of programs developed for LEP students should be to reduce or eliminate their English language deficit. These views are not shared by as many early-or late-exit teachers. Contrary to current research, more than two-thirds of immersion strategy and late-exit teachers feel that LEP students learn English better if their mistakes are detected and corrected as early as possible. Fewer early-exit teachers agree with this statement. Approximately half of all teachers in each program feel that learning content area knowledge in two languages more than doubles the learning effort for a LEP student. These program differences are consistent with the instructional model for each program.

Table 105

Percentage of Teachers Choosing Frequency with which Spanish Should be used for Instruction by Teachers, by Program and Grade

Grade	Frequency	Immersion %	Program Early-Exit %	Late-Exit %
K	Regularly for formal instruction	3.1	62.2	71.0
	Regularly for support and explanation	0.0	20.5	22.8
	Occasionally, to clarify	85.6	17.3	0.0
	Only as a last resort	11.2	0.0	0.0
	Never	0.0	0.0	0.0
	Other	0.0	0.0	6.2
1	Regularly for formal instruction	1.9	36.6	81.5
	Regularly for support and explanation	12.8	18.9	11.0
	Occasionally, to clarify	61.0	36.7	7.4
	Only as a last resort	13.8	0.0	0.0
	Never	7.1	4.4	0.0
	Other	3.4	3.4	0.0
2	Regularly for formal instruction	4.4	19.3	83.6
	Regularly for support and explanation	11.6	17.8	7.7
	Occasionally, to clarify	56.3	61.8	3.2
	Only as a last resort	15.5	1.1	0.0
	Never	4.1	0.0	0.0
	Other	8.2	0.0	5.5
3	Regularly for formal instruction	4.8	19.6	55.9
	Regularly for support and explanation	14.0	42.7	26.3
	Occasionally, to clarify	71.5	19.9	12.9
	Only as a last resort	9.7	13.9	0.0
	Never	0.0	4.0	0.0
	Other	0.0	0.0	5.0

Table 105 (continued)

Grade	Frequency	Immersion %	Program Early-Exit %	Late-Exit %
4	Regularly for formal instruction	0.0	13.2	51.4
	Regularly for support and explanation	0.0	14.2	24.2
	Occasionally, to clarify	76.6	57.8	19.3
	Only as a last resort	23.4	7.4	2.6
	Never	0.0	7.4	2.6
	Other	0.0	0.0	0.0
5	Regularly for formal instruction	*	*	42.8
	Regularly for support and explanation	*	*	39.6
	Occasionally, to clarify	*	*	11.0
	Only as a last resort	*	*	6.6
	Never	*	*	0.0
	Other	*	*	0.0
6	Regularly for formal instruction	*	*	65.1
	Regularly for support and explanation	*	*	12.7
	Occasionally, to clarify	*	*	11.1
	Only as a last resort	*	*	4.7
	Never	*	*	0.0
	Other	*	*	6.4
	Weighted Number of Responses	177	180	176

Table 106

Percentage of Teachers Choosing Frequency with which Spanish Should be used for Social Interaction with Students, by Program and Grade

Grade	Frequency	Immersion %	Program Early-Exit %	Late-Exit %
K	Regularly	9.4	62.0	90.0
	Occasionally	59.3	34.4	5.9
	Only as a last resort	13.2	3.6	0.0
	Never	14.4	0.0	0.0
	Other	3.8	0.0	4.1
1	Regularly	12.9	42.6	92.6
	Occasionally	46.6	46.1	7.4
	Only as a last resort	33.6	4.6	0.0
	Never	3.3	4.5	0.0
	Other	3.7	2.2	0.0
2	Regularly	19.0	28.7	83.6
	Occasionally	51.5	52.5	13.2
	Only as a last resort	13.4	9.4	3.2
	Never	12.1	0.0	0.0
	Other	4.0	9.5	0.0
3	Regularly	19.6	41.4	68.0
	Occasionally	36.9	36.9	22.1
	Only as a last resort	25.2	14.1	3.0
	Never	3.8	7.9	0.0
	Other	14.5	0.0	6.9

ERIC Full Text Provided by ERIC

Table 106 (continued)

Grade	Frequency	Immersion %	Program Early-Exit %	Late-Exit %
4	Regularly	0.0	14.2	57.4
	Occasionally	45.8	35.6	36.6
	Only as a last resort	36.2	35.4	2.6
	Never	18.1	14.8	0.0
	Other	0.0	0.0	3.4
5	Regularly	*	*	67.7
	Occasionally	*	*	29.6
	Only as a last resort	*	*	0.0
	Never	*	*	2.7
	Other	*	*	0.0
6	Regularly	*	*	68.4
	Occasionally	*	*	17.4
	Only as a last resort	*	*	0.0
	Never	*	*	4.7
	Other	*	*	9.5
	Weighted Number of Responses	181	180	178

ERIC

Table 107

Percentage of Teachers Choosing View of Spanish Use Among Students During Instruction, by Program and Grade

Grade	View	Program Immersion %	 Early-Exit %	 Late-Exit %
K	Should be discouraged	20.7	1.8	11.8
	Should be encouraged	6.3	42.2	58.1
	Should be accepted	73.1	56.0	30.1
	Other	0.0	0.0	0.0
1	Should be discouraged	16.9	7.8	3.0
	Should be encouraged	1.8	19.6	53.5
	Should be accepted	76.2	70.4	36.8
	Other	5.1	2.2	6.7
2	Should be discouraged	20.7	9.5	0.0
	Should be encouraged	4.4	15.5	67.6
	Should be accepted	70.9	72.6	30.1
	Other	4.0	2.4	2.4
3	Should be discouraged	7.6	21.6	5.9
	Should be encouraged	4.1	25.3	53.1
	Should be accepted	88.3	43.4	36.0
	Other	0.0	9.8	5.0
4	Should be discouraged	0.0	50.1	7.7
	Should be encouraged	0.0	0.0	26.8
	Should be accepted	81.9	42.5	65.6
	Other	18.1	7.4	0.0

Table 107 (continued)

Grade	View	Immersion %	Program Early-Exit %	Late-Exit %
5	Should be discouraged	*	*	3.6
	Should be encouraged	*	*	39.5
	Should be accepted	*	*	54.2
	Other	*	*	2.7
6	Should be discouraged	*	*	0.0
	Should be encouraged	*	*	46.5
	Should be accepted	*	*	37.7
	Other	*	*	15.8
	Weighted Number of Responses	181	181	178

ERIC Full Text Provided by ERIC

Table 108

Percentage of Teachers Choosing View of Spanish Use Among LEP Students for Social Interaction, by Program and Grade

Grade	View	Program Immersion %	 Early-Exit %	 Late-Exit %
K	Should be discouraged	0.0	0.0	11.8
	Should be encouraged	16.9	45.7	66.3
	Should be accepted	56.3	47.2	11.8
	Other	26.8	7.1	10.0
1	Should be discouraged	16.4	6.6	4.4
	Should be encouraged	5.5	15.1	48.8
	Should be accepted	71.2	75.2	40.2
	Other	6.9	3.2	6.7
2	Should be discouraged	8.6	8.5	0.0
	Should be encouraged	6.5	13.0	48.9
	Should be accepted	84.9	76.1	51.1
	Other	0.0	2.4	0.0
3	Should be discouraged	3.8	14.5	5.8
	Should be encouraged	0.0	18.3	55.7
	Should be accepted	96.2	65.1	30.7
	Other	0.0	2.1	7.8
4	Should be discouraged	18.1	29.5	5.1
	Should be encouraged	18.1	0.0	26.9
	Should be accepted	63.9	70.5	68.0
	Other	0.0	0.0	0.0

Table 108 (continued)

Grade	View	Immersion %	Program Early-Exit %	Late-Exit %
5	Should be discouraged	*	*	0.0
	Should be encouraged	*	*	46.3
	Should be accepted	*	*	51.0
	Other	*	*	2.7
6	Should be discouraged	*	*	0.0
	Should be encouraged	*	*	42.2
	Should be accepted	*	*	46.2
	Other	*	*	11.6
	Weighted Number of Responses	181	180	177

Table 109

Percentage of Teachers Identifying Goals as Program Goals, by Program and Grade

Grade	Goal	Immersion %	Program Early-Exit %	Late-Exit %
K	Develop English	96.9	[illegible].4	82.2
	Develop Spanish	0.0	8[illegible]9	82.0
	Master content areas	96.9	100[illegible]	100.0
	Success in English-only program	88.8	63.[illegible]	30.1
	Social integration	96.9	100.0	88.2
	Maintenance of Hispanic culture	79.5	98.2	88.2
1	Develop English	100.0	97.8	93.5
	Develop Spanish	4.1	78.5	95.6
	Master content areas	94.7	100.0	100.0
	Success in English-only program	98.2	69.2	30.2
	Social integration	95.9	95.3	95.6
	Maintenance of Hispanic culture	85.4	84.2	91.2
2	Develop English	100.0	98.8	92.9
	Develop Spanish	12.8	58.6	100.0
	Master content areas	100.0	100.0	100.0
	Success in English-only program	95.1	64.1	31.6
	Social integration	97.8	92.8	97.6
	Maintenance of Hispanic culture	73.2	82.9	92.9
3	Develop English.	100.0	100.0	98.5
	Develop Spanish	4.8	65.2	79.1
	Master content areas	100.0	100.0	95.7
	Success in English-only program	95.2	53.6	68.4
	Social integration	100.0	100.0	92.7
	Maintenance of Hispanic culture	83.7	100.0	86.9

Table 109 (continued)

Grade	Goal	Immersion %	Program Early-Exit %	Late-Exit %
4	Develop English	100.0	100.0	100.0
	Develop Spanish	0.0	43.3	90.6
	Master content areas	100.0	100.0	100.0
	Success in English-only program	77.1	93.4	49.3
	Social integration	100.0	100.0	100.0
	Maintenance of Hispanic culture	81.9	86.8	85.7
5	Develop English	*	*	100.0
	Develop Spanish	*	*	84.9
	Master content areas	*	*	100.0
	Success in English-only program	*	*	57.8
	Social integration	*	*	100.0
	Maintenance of Hispanic culture	*	*	93.7
6	Develop English	*	*	100.0
	Develop Spanish	*	*	88.9
	Master content areas	*	*	100.0
	Success in English-only program	*	*	67.5
	Social integration	*	*	100.0
	Maintenance of Hispanic culture	*	*	93.3
	Weighted Number of Responses	179 - 181	178 - 180	174 - 178

Table 110

Percentage of Teachers Agreeing with Statement (Strongly or Somewhat), by Program

Statement	Program: Immersion %	Early-Exit %	Late-Exit %
LM-LEP students need to develop skills in their primary language similar to the skills they develop in English	61.9	86.1	97.7
LM-LEP students who have learned good conversational English are ready for content area instruction given entirely in English	82.0	69.9	56.9
LM-LEP students learn to read English best if they are first taught to read in their primary language	40.3	79.1	95.7
If LM-LEP students are taught content areas in their primary language at home, the school does not need to teach these content areas in that language	49.4	24.7	6.2
A teacher best uses a LM-LEP student's primary language to support primary instruction given in English, rather than using it as a primary language of instruction	88.4	57.2	31.4
Learning content area knowledge in two language more than doubles the learning effort for a LM-LEP student	58.0	50.2	49.1
LM-LEP students learn English better if their mistakes are detected and corrected as early as possible	67.6	57.8	71.9
LM-LEP students who have a good control over oral English learn to read in English more easily than those whose oral English is weak	86.4	87.2	89.1

Table 110 (continued)

Statement	Immersion %	Program Early-Exit %	Late-Exit %
IM-LEP students learn English best by using it to communicate inside and outside the classroom	95.7	90.2	94.7
IM-LEP students' ability to speak English develops more slowly than their ability to comprehend English	77.9	78.8	68.3
IM-LEP students are helped by having content area lessons, given in English, concurrently translated into their primary language	30.4	47.4	32.9
How well IM-LEP students know their primary language is important in deciding how or what to teach them in school	68.6	80.9	89.5
The main purpose of programs developed for IM-LEP students should be to reduce or eliminate their English language deficit	91.6	82.8	69.5
Range of Weighted Number of responses	175-181	176-181	170-179

Conclusion

With a few exceptions, teachers in the immersion strategy, early-exit and late-exit programs tend to be similar with respect to characteristics, attitudes, and behaviors which do not relate to the program models. Differences tend to occur on those characteristics, attitudes, and behaviors which derive from the different underlying program philosophies and models.

Late-exit teachers generally are of a different ethnic background (resembling that of their students) than immersion strategy and early-exit teachers, who are quite similar in ethnic background to each other (and somewhat different from their students). A higher percentage of late-exit teachers tend to have completed more schooling and have had more ESL training than teachers in the other programs. Fewer immersion strategy and early-exit teachers have specialized training as grade level increases. Teachers in all programs are sufficiently proficient in English to teach in English. While late-exit teachers are sufficiently proficient to teach in Spanish, early-exit teachers are not. The latter raises the question of fidelity/quality of treatment for the early-exit programs.

Consistent with the program models, immersion strategy teachers begin English reading instruction without considering children's Spanish literacy, while early-exit and late-exit teachers report using children's Spanish literacy as a criterion to the introduction of English reading instruction. Immersion strategy teachers tend to devote more time to English language arts than early-exit or late-exit teachers, although teachers in these two programs report increasing the proportion of time devoted to English language arts in the higher grades. However, early-exit and late-exit teachers provide Spanish language arts instruction, so that the total amount of instructional time spent on language arts, English and/or Spanish, is comparable across the three programs. Immersion strategy teachers instruct primarily in English while early-exit teachers use varying amounts of

Spanish, and late-exit teachers use primarily Spanish for instruction, with more English in the higher grades.

Teachers in all three programs tend to account for the same amount of total instructional time per day, allocating approximately the same amount of instruction for each content area. Late-exit teachers assign and grade homework with greater frequency than immersion strategy or early-exit teachers. Student misbehavior generally interferes with instruction to a small extent or not at all within each program. Average class sizes are similar across the three programs, and the proportions of LEP and FEP students in classrooms is comparable across programs within each grade. Teachers in all programs receive very few services from resource teachers or classroom aides.

Teachers in the three programs make similar use of written programs of instruction for English skills, although late-exit teachers tend to use different types of materials from immersion strategy and early-exit teachers.

Immersion strategy, early-exit and late-exit teachers have similar priorities regarding essential support services, and their ratings of program strengths and weaknesses also are not very different. Except for the development of Spanish and success in an English-only program, teachers in all programs tend to agree about program goals. Teacher attitudes are moderately consistent with the three program models.

Administrative Context

Introduction

This chapter provides information on the administrative context of the programs in the study. Instructional programs are affected by administrative practices, school and district policies, and conditions, whether or not these are designed with the program model in mind. Superintendents, principals, and project administrators make daily decisions which can support or impede project objectives. In addition, unique situations at the school or district level can produce changes in implementation of a particular program model. With these possibilities in mind, interviews were conducted with project administrators, site administrators (principals), and district administrators. The following sections present the results of these interviews and discuss administrative differences between programs if such differences appear to exist.

Project Administrators

Of the nine districts included in this study, four offer both immersion strategy and early-exit programs, three offer only late-exit programs, one offers only an immersion strategy program, and one offers only an early-exit program. Only one district administers the immersion strategy and early-exit programs separately, so that, in most cases, whoever administers the early-exit program also administers the immersion strategy program. The description that follows separates the policies for each program by duplicating the data provided by the administrators of both immersion strategy and early-exit programs and weighting appropriately.

ERIC

What are the background characteristics of project administrators and do they differ by program?

Ten project directors administered the 13 programs included in this study. Of these, two administered an immersion strategy program, two administered an early-exit program, three administered a late-exit program, and three administered a joint program which included both immersion strategy and early-exit.

Differences among the background characteristics of project directors by program were evident. While across programs slightly less than two-thirds of the administrators were Hispanic (59.4%) and slightly more than one-third were white (40.6%), program differences were evident (see Table 111). Two-thirds of the immersion strategy program administrators were white, while early-exit administrators tended to be almost equally divided between white (54.8%) and Hispanic (45.2%). In marked contrast, all of the late-exit project directors were Hispanic. Across programs, four-fifths of the project directors spoke Spanish, two-fifths as their primary language. Slightly less than three-fourths of the immersion strategy and early-exit project directors spoke Spanish (71.3% and 68.5%, respectively); approximately one-tenth of the immersion strategy and early-exit did so as their primary language (13.3% and 14.6%, respectively). In contrast, all of the late-exit project directors spoke Spanish as their primary language. Across programs, project administrators were almost equally divided between those who had earned a master's degree (53.8%) and those who had earned a doctorate (46.2%). While approximately one-third of the immersion strategy and early-exit directors had completed their doctorates (30.1% and 31.5%, respectively), no late-exit administrator had done so. Overall, almost two-thirds of the administrators had a bilingual teaching credential (63.6%) and almost half had an ESL credential (52.2%). Approximately half of the immersion strategy and early-exit teachers had a bilingual teaching credential (50.5% and 53.9%, respectively) and half had an ESL credential (50.5% and 45.7%, respectively). Additionally, almost one-third of the immersion strategy and early-exit project directors were enrolled in a bilingual credentialing program (28.8% and 31.6%, respectively) and/or an ESL certification program (28.8% and 39.8%, respectively). In contrast, over four-fifths of the late-exit project directors had a bilingual teaching credential and nearly two-thirds had an ESL credential (86.5% and 60.1%, respectively). No late-exit administrator was enrolled in a bilingual or ESL credentialing program. Overall, project administrators had an average of 5.6 years of bilingual/ESL teaching experience.

Table 111

Project Administrator Characteristics, by Program

Characteristics	Type of Program: Immersion %	Early-Exit %	Late-Exit %	Overall %
Ethnicity:				
Hispanic	33.7	45.2	100.0	59.4
White	66.3	54.8	0.0	40.6
Primary language:				
Spanish	13.3	14.6	100.0	42.6
English	86.7	85.4	0.0	57.4
Percentage who speak Spanish	71.3	68.5	100.0	79.9
Academic degrees:				
Master's	69.9	68.5	100.0	53.8
Doctorate	30.1	31.5	0.0	46.2
Certified bilingual teacher	50.5	53.9	86.5	63.6
Certificate in ESL	50.5	45.7	60.1	52.2
Enrolled in bilingual credential program	28.8	31.6	0.0	20.1
Enrolled in ESL certificate program	28.8	39.8	0.0	22.7
Mean years of bilingual/ESL teaching experience	4.3 yrs.	5.3 yrs.	7.3 yrs.	5.6 yrs.

When asked about having had special training to work with second language learners, all late-exit project directors reported that they had training in ESL, second language acquisition, first language development, multicultural education, bilingual education, and language assessment (see Table 112). In contrast, as few as one-half of immersion strategy and early-exit administrators had specialized training in some of these areas.

Table 112

Percentage of Project Administrators Reporting Training, by Program

	Type of Program			
Topic	Immersion	Early-Exit	Late-Exit	Overall
ESL	58.0	53.9	100.0	70.7
Second language acquisition	71.3	68.5	100.0	79.9
First language development	71.3	68.5	100.0	79.9
Multicultural education	86.7	85.4	100.0	90.7
Bilingual education	79.2	85.4	100.0	88.1
Language assessment	100.0	100.0	100.0	100.0

In sum, as a group, project directors tended to be Hispanic, speak Spanish, have an advanced educational degree, and have had training and experience as bilingual and/or ESL teachers. However, there are strong differences by program. Immersion strategy program administrators tended to be white, speak Spanish as their second language, and have a master's degree or better; only one-half have been trained as bilingual or ESL teachers. Among the three programs, they had the least amount of bilingual/ESL teaching experience (X = 4.3 years). While early-exit project directors were as likely to be white as Hispanic, most spoke Spanish as their second language, had a master's degree or better, and about one-half had a bilingual or ESL teaching credential. The late-exit administrators were all Hispanic, spoke Spanish as their primary language, and had a master's degree. Most of them had certification as a bilingual and/or ESL teacher, and they had the greatest amount of bilingual/ESL teaching experience (X = 7.3 years).

How do project administrators believe Spanish should be used in the classroom?

The extent to which Spanish is used in the classroom, and the ways in which it is used, represent an important distinction between the three programs under consideration. To some extent, the attitudes of the administrators reflect the theoretical differences which distinguish the separate programs. For example, when asked whether teachers should use Spanish in the classroom to teach or whether it should be used regularly in social interactions with their students, between one-half and two-thirds of the administrators in the immersion strategy and early-exit programs thought Spanish should be used regularly for instruction and in social interactions with their students (see Tables 113 and 114). In contrast, over four-fifths of the administrators in the late-exit program felt that Spanish should be used regularly for instruction and social interaction with their students. When asked their opinions about whether LEP students should use Spanish during instruction or in social interaction, surprisingly, approximately one-quarter of all project administrators felt that the use of Spanish by students during instruction should be discouraged (immersion strategy, 20.4%; early-exit, 30.6%; and late-exit, 26.4%) (see Table 115). This unexpected response is inconsistent with the underlying tenet of immersion strategy programs not to "force" second language learners to produce the target language until they are ready to produce it on their own. Similarly, the development of any language requires that it be used in a variety of contexts. In bilingual programs this implies that the primary language should be used for instruction. Yet, in both the early- and late-exit programs some administrators felt that the use of Spanish for instruction by students should be discouraged. These responses are all inconsistent with each of the three programs. While more immersion strategy and early-exit administrators were accepting of LEP students using Spanish for social interaction (86.7% and 85.4%, respectively), approximately one-fourth of the late-exit administrators felt that LEP students should be discouraged from using Spanish for

ERIC

social interaction (see Table 116). This view by the late-exit program administrators is inconsistent with their program model.

Table 113

Project Administrators' Attitudes Toward Use of Spanish to Instruct LEP Students, by Program

	Type of Program			
Response	Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be used regularly	56.6	62.1	86.5	68.3
Should be used only occasionally	30.1	0.0	13.5	15.0
Should be used only as a last resort	0.0	0.0	0.0	0.0
Should never be used	0.0	0.0	0.0	0.0
Other	13.3	37.9	0.0	16.7

Table 114

Project Administrators' Attitudes Toward Use of Spanish for Social Interactions with LEP Students, by Program

	Type of Program			
Response	Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be used regularly	56.6	62.1	86.5	68.3
Should be used only occasionally	0.0	0.0	13.5	4.5
Should be used only as a last resort	30.1	0.0	0.0	10.5
Should never be used	0.0	0.0	0.0	0.0
Other	13.3	37.9	0.0	16.7

Table 115

Project Administrators' Attitudes Toward LEP Students' Use of Spanish During Instruction, by Program

Response	Type of Program: Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be discouraged	20.4	30.6	26.4	25.6
Should be encouraged	0.0	0.0	60.1	20.0
Neither encouraged nor discouraged, but accepted	36.2	31.6	13.5	27.2
Other	43.4	37.9	0.0	27.2

Table 116

Project Administrators' Attitudes Toward LEP Students' Use of Spanish In Classroom for Social Interaction, by Program

Response	Type of Program: Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be discouraged	0.0	0.0	26.4	8.8
Should be encouraged	0.0	0.0	0.0	0.0
Neither encouraged nor discouraged, but accepted	86.7	85.4	73.6	81.9
Other	13.3	14.6	0.0	9.3

<u>How well are programs articulated? Does this differ by program?</u>

While all programs have developed written curriculum objectives for most content areas, early- and late-exit programs are more likely to have such plans in Spanish, while the immersion strategy programs generally have such plans for language arts only in English (see Table 117). This difference would appear to be consistent with program goals for the immersion strategy programs, which place heavier emphasis on English. Fewer than one-half of the late-exit project administrators report having written curriculum objectives for English reading, writing and speaking. All project administrators reported that teachers were required to follow the written curricula, and coordination took place in a variety of ways, including (a) having the principal review the weekly lesson plans, (b) reviewing objectives at regular planning meetings between principals and resource teachers and between project directors and resource teachers, and (c) having project staff visit classrooms to observe classes (see Table 118).

Table 117

Percentage of Administrators Reporting Use of Written Program Objectives, by Program

Subject	Type of Program Immersion %	 Early-Exit %	 Late-Exit %
English reading	100.0	100.0	39.9
English writing	100.0	100.0	13.5
English speaking	71.3	68.5	39.9
Spanish reading	0.0	100.0	100.0
Spanish writing	0.0	68.5	73.6
Spanish speaking	0.0	100.0	100.0
Math	100.0	100.0	100.0
Science	100.0	100.0	100.0
Social Studies	86.7	85.4	100.0

Table 118

Percentage of Administrators Reporting Methods of Monitoring Use of Curriculum Plans, by Program

Method of Monitoring	Type of Program: Immersion %	Early-Exit %	Late-Exit %
Principal (or surrogate) reviews lesson plans weekly	100.0	100.0	100.0
Regular planning meetings between principal and/or project resource teacher to review objectives to be taught by staff during a particular time period, etc.	100.0	100.0	100.0
Regular planning meetings between resource teachers and project director to review implementation of instructional objectives by project staff	100.0	100.0	100.0
Direct classroom observation by central office project staff	100.0	100.0	100.0

What are the goals of each program and how do they differ?

One of the most marked differences among programs appeared when project administrators were asked what they understood to be the goals of their programs for LEP students (see Table 119). Agreement among project administrators was unanimous regarding the development of English and the successful mastery of content areas as program goals. While all immersion strategy and early-exit project administrators reported that successful participation in an English-only program was an objective, only one-third of the late-exit administrators felt that this was a program goal. In contrast, all late-exit administrators indicated that development of Spanish was a program goal, while only two-thirds of the early-exit and none of the immersion strategy administrators felt that this was a program objective. These results are consistent with each program model.

Table 119

Percentage of Administrators Naming Program Goals, by Program

Program Goals	Immersion %	Early-Exit %	Late-Exit %
Develop English	100.0	100.0	100.0
Develop Spanish	0.0	68.5	100.0
Successful mastery of content areas	100.0	100.0	10C.0
Successful participation in an English-only program	100.0	100.0	33.8

What are the entry and exit criteria for each program?

Not surprisingly, all programs share more or less the same criteria for assigning children to classes for LM-LEP students (see Table 120). All programs test children's English speaking and comprehension skills for entry and exit. While nearly all of the programs also test children's knowledge of reading and writing in English for exit, all of the late-exit, yet only two-thirds of the early-exit and one-fourth of the immersion strategy programs do so for entry. Roughly two-thirds of the late-exit programs assess Spanish oral language skills for entry, but do not do so for exit purposes. For all practical purposes, oral Spanish skills are not assessed by either immersion strategy or early-exit programs. None of the three programs assess Spanish reading or writing skills for either entry or exit. Teacher judgment, school or district committee recommendations, and parent approval all appear to play a greater role in immersion strategy and early-exit programs than in late-exit programs for determining entry into LM-LEP classes or exit to mainstream, English-only classes, but these are rarely the deciding factor. In most programs, assessment is a complex process taking into account a number of the factors cited above.

Table 120

Percentage of Administrators Reporting Entry and Exit Criteria, by Program

Criteria	Type of Program		
	Immersion %	Early-Exit %	Late-Exit %
Oral English			
Entry criterion	100.0	100.0	100.0
Exit criterion	100.0	100.0	100.0
English reading/writing			
Entry criterion	28.8	63.0	100.0
Exit criterion	100.0	100.0	86.5
Oral Spanish			
Entry criterion	0.0	8.2	60.1
Exit criterion	0.0	8.2	0.0
Spanish reading/writing			
Entry criterion	0.0	0.0	0.0
Exit criterion	0.0	0.0	0.0
Teacher judgment			
Entry criterion	13.3	22.8	13.5
Exit criterion	50.9	46.1	39.9
School/District committee recommendation			
Entry criterion	49.2	62.1	26.4
Exit criterion	92.5	100.0	26.4
Parent approval			
Entry criterion	62.5	76.7	26.4
Exit criterion	62.5	76.7	26.4

<u>Are teachers provided with special training, and does this differ by program?</u>

Project directors in all three programs reported that teachers are provided with inservice on ESL methodology and second language acquisition. All project administrators of immersion strategy and early-exit programs, yet only three-quarters of late-exit administrators (73.6%) reported that teachers receive training in language assessment procedures. Teachers in early- and late-exit programs are provided with inservice on first language acquisition and bilingual education. Slightly less than one-half of the immersion strategy project directors reported that teachers receive training on first language acquisition (46.9%) and about four-fifths reported that teachers receive training

ERIC

on bilingual education (78.0%). With the exception of training in first language acquisition which fewer immersion strategy districts provide to teachers, neither the provision of such training nor the choice of topics seems to be a function of the type of program offered, but seems to be more a matter of district policy.

It is interesting to note that six of the nine districts in the study changed their entry and/or exit criteria within the past seven years. These changes resulted from changes in state laws and court-ordered desegregation, as well as lack of administrative support for the program and concern over the length of time students remain in bilingual programs.

Summary

In sum, project administrators of the study's programs for LM-LEP students generally are knowledgeable of the needs of limited-English proficient children and have had substantial experience in working with them. Across programs project directors tend to be Hispanic and bilingual. However, by program, immersion strategy project administrators tended to be white with Spanish as their second language, while administrators in early-exit programs tended to be evenly divided between white and Hispanic and also had Spanish as their second language. All late-exit administrators were Hispanic and native Spanish speakers. More late-exit project administrators had extensive training in the areas of bilingual education, language development and acquisition than immersion strategy or early-exit administrators. In many cases, their attitudes toward the education of children with limited proficiency in English reflected the underlying theories and policies of their respective programs. Such differences are particularly clear regarding the use of Spanish in the classroom and the goals of programs for LM-LEP students.

ERIC

Site Administrators

A total of forty-five site administrators (principals) replied to the site administrator questionnaire, fifteen from school sites having immersion strategy only programs, five from sites with both immersion strategy and early-exit programs, eleven from sites with only early-exit programs, and fourteen from sites with only late-exit programs. Each school within a district received equal weight regardless of whether it had one program or two (see Appendix A). The following section presents the results of these questionnaires and compares results across programs.

What are the background characteristics of site administrators? Does this differ by program?

Of the site administrators responding, across programs, more than one-half were white (57.5%) and four-fifths had English as their primary language (82.7%), with about one-half reporting that they could speak Spanish (55.2%) (see Table 121). Over four-fifths had completed a master's degree (89.1%); however, few had a bilingual or ESL teaching credential (20.0%), or had taught bilingual or ESL classes (21.3% and 37.3%, respectively), nor had they taught in early-exit or late-exit bilingual classes. The majority of principals have had special training in program administration (94.4%) and staff development (94.4%) (see Table 122). Most also reported having completed special training or coursework in ESL, second language acquisition, first language development, multicultural education, bilingual education and program design. However, large differences by program are evident.

Proportionately more principals in late-exit and early-exit programs were Hispanic than in the immersion strategy program (43.9% and 29.7% vs. 3.4%). While one-third of the principals in the late-exit program spoke Spanish as their first language, a considerably smaller proportion of the principals in the early-exit and immersion strategy programs did so (12.4% and 3.7%, respectively). Early-exit principals tended to have completed more advanced degrees than their immersion strategy or late-exit counterparts. Generally, only small differences in type of training are found between the late-exit and early-exit principals. Early- and late-exit principals had more experience and training than did principals in the immersion strategy program. Proportionately more of the principals in the late-exit program than those in immersion strategy or early-exit programs had obtained a bilingual and/or ESL teaching credential, had experience in bilingual classrooms, or had taught in early- and late-exit classrooms. Moreover, late-exit administrators reported having had more specialized training or coursework in first language development, multicultural education, bilingual education, program administration, and staff development than principals in the early-exit or immersion strategy programs.

ERIC

Table 121

Percentage of Site Administrators (Principals) Having Characteristics

Characteristics	Type of Program Immersion %	 Early-Exit %	 Late-Exit %	Overall %
Ethnicity: Hispanic	3.4	29.7	43.9	27.1
White	72.4	62.8	39.0	57.5
Black	16.9	2.4	17.2	11.4
Primary language:				
Spanish	3.7	12.4	35.1	17.8
English	96.6	87.9	65.0	82.7
Speak Spanish:	33.5	72.4	52.4	55.2
Academic degrees:				
Bachelor's	0.0	0.0	8.6	2.9
Master's/Specialist	100.0	79.4	91.4	89.1
Doctorate	0.0	20.6	0.0	8.0
Bilingual or ESL certification:	0.0	21.4	35.1	20.0
Experience in bilingual classrooms:				
None	96.6	82.4	58.7	78.7
Less than three years	3.4	10.3	4.6	6.5
Three to five years	0.0	7.3	22.6	10.2
More than five years	0.0	0.0	14.1	4.6
Have taught in early-exit bilingual classes:	0.0	17.6	22.6	14.3
Have taught in late-exit bilingual classes:	0.0	0.0	18.7	6.1
Experience in ESL classrooms:				
None	93.1	48.0	54.1	62.7
Less than three years	0.0	29.7	18.2	17.6
Three to five years	3.4	22.4	0.0	9.8
More than five years	3.4	0.0	27.7	9.9

Table 122

Percentage of Principals Reporting Training or Coursework

Type of Training	Type of Program: Immersion %	Early-Exit %	Late-Exit %	Overall %
ESL	36.8	76.4	64.1	61.2
Second language acquisition	25.7	76.4	63.8	58.0
First language development	48.6	76.4	100.0	76.8
Multicultural education	48.1	89.5	91.0	77.9
Bilingual education	31.9	76.4	82.0	65.6
Program administration	93.3	90.5	100.0	94.4
Staff development	93.3	90.5	100.0	94.4
Program design	71.3	79.0	77.4	76.3
Other	20.4	29.2	50.3	35.1

How do site administrators believe English and Spanish should be used in the classrooms? Does this differ by program?

As noted earlier, the proportion of English and Spanish used in the classroom, as well as how they are used, differentiates the three instructional programs under study. Much has been written regarding the importance of the school site administrator's role as instructional leader. Consequently, the site administrator's attitudes towards the role of English and Spanish in the classroom would likely influence how teachers would use each language.

As expected, the majority of site administrators across programs reported that English should be used regularly for instruction and when interacting socially with students (78.3% and 77.1%, respectively) (see Tables 123 and 124). However, program differences were evident, and they were consistent with the instructional model of each program. While almost all of the immersion strategy site administrators felt that English should be used regularly for instruction with LM-LEP students, fewer early-exit and late-exit administrators felt the same way (92.8% vs. 82.8% and 61.0%). Proportionately more immersion strategy and early-exit administrators than late-exit administrators

ERIC

felt that English should be used regularly when school personnel interact socially with LEP students (86.1% and 84.5% vs. 61.0%).

Site administrators were also asked their opinions as to whether LEP students should use Spanish during instruction and/or for social interaction. Overall, about one-half of the site administrators indicated that the use of Spanish for instruction should neither be encouraged nor discouraged, but accepted (see Table 125). About one-third felt that students should be encouraged to use Spanish during instruction. As before, program differences reflect the instructional model of each program. Proportionately more immersion strategy and early-exit site administrators than late-exit administrators felt that they should discourage students from using Spanish during instruction (24.8% and 14.0% vs. 0.0%). While considerably more immersion strategy and early-exit site administrators were accepting of students' use of Spanish for social interaction than late-exit administrators (82.8% and 64.2% vs. 43.4%), proportionately more of them felt that students should be discouraged from using Spanish for this purpose than did late-exit administrators (17.2% and 9.5% vs. 0.0%) (see Table 126).

In sum, the use of English and Spanish by instructional staff and students reflected the instructional model of the programs [illegible] site administrators were supervising.

ERIC

Table 123

Site Administrators' Attitudes Toward Use of English to Instruct LEP Students, by Program

	Type of Program			
Response	Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be used regularly	92.8	82.8	61.0	78.3
Should be used only occasionally	7.2	17.2	17.2	14.4
Should be used only as a last resort	0.0	0.0	0.0	0.0
Should never be used	0.0	0.0	0.0	0.0
Other	0.0	0.0	21.8	7.3

Table 124

Site Administrators' Attitudes Toward Use of English for Social Interactions with LEP Students, by Program

	Type of Program			
Response	Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be used regularly	86.1	84.5	61.0	77.1
Should be used only occasionally	13.9	15.5	21.6	17.1
Should be used only as a last resort	0.0	0.0	0.0	0.0
Should never be used	0.0	0.0	0.0	0.0
Other	0.0	0.0	17.4	5.8

Table 125

Site Administrators' Attitudes Toward LEP Students' Use of Spanish During Instruction, by Program

Response	Type of Program: Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be discouraged	24.8	14.0	0.0	12.3
Should be encouraged	13.9	26.7	56.6	33.1
Neither encouraged nor discouraged, but accepted	57.9	59.3	30.4	49.3
Other	3.4	0.0	13.0	5.3

Table 126

Site Administrators' Attitudes Toward LEP Students' Use of Spanish In Classroom for Social Interaction, by Program

Response	Type of Program: Immersion %	Early-Exit %	Late-Exit %	Overall %
Should be discouraged	17.2	9.5	0.0	8.5
Should be encouraged	0.0	26.3	43.6	24.8
Neither encouraged nor discouraged, but accepted	82.8	64.2	43.4	62.4
Other	0.0	0.0	13.0	4.3

What do site administrators understand their program goals to be? Does this differ by program?

While there is general agreement among principals across programs as to the development of English as a program goal, there are marked program differences with respect to other goals. All immersion strategy and early-exit principals and most late-exit principals reported that the development of English was a program goal (see Table 127). While all or most immersion strategy principals indicated that mastering content areas, succeeding in English-only classrooms, promoting social integration, and maintaining Hispanic culture were program goals, noticeably fewer late-exit principals shared these views (31.4% to 78.4%). Considerably fewer late-exit administrators thought that one of their goals was to prepare students for successful participation in English-only classrooms (33.0%) or to promote social integration (31.4%). Reflecting program philosophies, the majority of early- and late-exit principals held that developing Spanish language skills was a program objective (97.7% and 78.4%, respectively).

Table 127

Percentage of Site Administrators Selecting Program Goals

	Type of Program		
Stated Program Goals	Immersion %	Early-Exit %	Late-Exit %
Develop English	100.0	100.0	87.0
Develop Spanish	10.3	97.7	78.4
Master content	100.0	100.0	43.2
Success in English-only classrooms	100.0	89.5	33.0
Social integration	100.0	89.2	31.4
Maintain Hispanic culture	78.8	100.0	57.4

Summary

Proportionately more principals in the early- and late-exit bilingual programs tended to have backgrounds that were more similar to that of their students, had been trained to work with language minority children, and had more practical experience in teaching these children than did principals in immersion strategy programs.

Site administrators generally understood the theoretical and methodological concepts underlying their respective programs. Early-exit and immersion strategy administrators emphasized the importance of developing students' skills in English, while late-exit administrators also emphasized improving Spanish language skills. While principals across programs viewed the development of English as a program goal, there was greater agreement between immersion strategy and early-exit

ERIC

principals than among late-exit principals regarding other program goals, such as mastering content, succeeding in English-only classrooms, and social integration.

In sum, overall, principals were not as well prepared in terms of training and experience in meeting the needs of language minority children as were project directors. More importantly, principals' attitudes and opinions regarding the education of language minority children did not agree with those of their respective project administrators. It is possible that the principals' lack of training and experience in bilingual instructional programs accounts for their different responses.

School Context

For the forty-six schools surveyed, fourteen offer a late-exit program and the remainder offer either early-exit and/or immersion strategy programs. For some schools which have both immersion strategy and early-exit classrooms, only the immersion strategy classrooms are included in the study. Thus, for purposes of this study, sixteen schools offer only an immersion strategy program, thirteen schools offer only an early-exit program, and five schools offer both. The following section provides information on these schools and evaluates whether differences among schools result in systematic differences among programs.

What proportion of the student body is language-minority and limited-English proficient? Does this differ by program?

The review of the literature in Chapter I noted the evidence regarding the importance of having the opportunity to produce language as a necessary condition for acquiring language. Students provide one another with an important source of language learning opportunities through their interactions within and outside the classroom. The availability of native English-speaking models provides a markedly different language learning environment than a setting in which few or no native English-speakers are available.

Table 128 lists the average proportion of students that are language minority and limited-English-proficient by program. These data note that early-exit only sites or sites that have both an immersion strategy and an early-exit program have a higher language minority population (79.7% and 72.8%, respectively) than do either late-exit only or immersion strategy only sites (55.2% and 40.2%, respectively). Comparing the proportion of the student body that is Spanish-speaking language minority with the proportion of all language minority students at the participating school sites, it is clear that almost all of the language-minority students in the study sites are Spanish speaking.

A sizeable portion of the student body within each program is limited-English-proficient (see Table 128). Approximately one-quarter to one-half of the student population in each program is limited-English-proficient, with early-exit and immersion strategy/early-exit sites having the highest proportion of LEP students (48.1% and 48.4%, respectively) as compared to either immersion strategy only or late-exit only programs (28.7% and 29.8%, respectively). Consistent with the finding that the sites represent a largely Spanish-speaking language-minority population, the majority of the LEP students are Spanish speakers.

In sum, while each of the program sites had sizeable proportions of their student bodies that were language minority and LEP, when compared across programs, early-exit only and immersion strategy/early-exit sites contain a large majority of language minority students. From a language development point of view, this suggests the limited availability of native English speaking models for second language learners in these two types of program sites.

Table 128

Average Percentage Distributions of Spanish-Speaking Students, by Program Within School

	Type of Program			
	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
Percent of student body which is language minority (LM)	40.2	79.7	72.8	55.2
Percent of student body which is Spanish-speaking	34.4	76.5	68.2	50.5
Classification of student body:				
LEP	28.7	48.1	48.4	29.8
FEP	11.5	31.7	24.3	25.4
Classification of Spanish-speaking student body:				
LEP	24.7	46.8	44.2	26.8
FEP	9.7	29.7	24.0	23.7

How long have the LM-LEP students lived in this country? Does this differ by program?

School site administrators provided estimates of the residency histories of the LM-LEP students in the participating schools. These data are displayed by program in Table 129. An estimated one-half of the students at the schools with only early-exit programs are U.S.-born LM-LEP students, and one-half are foreign-born students. Schools that have programs that are immersion strategy only, immersion strategy/early-exit, and late-exit have only one-fifth to one-third of their students who are foreign-born students. Given this diversity of their respective student populations, the three programs are not comparable in terms of students' residency.

Table 129

Mean Percentages of LM-LEP Students With Residency, by Program

Residency	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
	Type of Program			
U.S. born	70.0	50.6	82.5	66.0
Foreign-born, in U.S. 1 year or less	10.3	9.4	2.0	8.5
Foreign-born, in U.S. 1-2 years	6.5	19.4	2.8	9.0
Foreign-born, in U.S. 2 years or more	13.3	24.2	13.0	16.6

What is the mobility rate of LEP students in the schools and how does it compare to the overall mobility rate of the school? Does this differ by program?

Continuity of instruction is critical to the learning process. Site administrators were asked to estimate the proportion of students who start school in September and are still in school in June. These data are presented in Table 130 by program. While there was no more than a 10% difference in the total school mobility rate between programs, differences in the mobility of their respective LEP students are striking. Immersion strategy LEP students are the most stable, with over four-fifths of these students remaining for the entire school year. Approximately three-fourths of the LEP students at the late-exit only and early-exit only program sites completed the entire school year (74.9% and 71.0%, respectively). LEP students in sites having both an immersion strategy and early-exit program have the highest mobility rate, with only three-fifths of the students completing the school year.

Table 130

Mean Percentages of Students Remaining for a Full School Year, by Program

	Type of Program			
Population	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
Total school	78.0	73.8	70.0	79.8
LEP students	83.8	71.0	59.8	74.9

<u>What is the socioeconomic level of the student population of the project schools? Are there differences across programs?</u>

The literature documenting the relationship of socioeconomic factors and student achievement is extensive. Socioeconomic factors can limit or expand the resources and learning opportunities for students. As with language proficiency, the relative mix of students from different socioeconomic levels within a school can either be a source for support and/or expansion of student learning, or it can constrain and minimize student learning. To this end, site administrators were asked to describe the background of their student body in terms of different socioeconomic factors.

Site administrators were asked to estimate the income levels of their students' families (see Table 131). While proportionately more of the student body in immersion strategy only program sites were judged to be of the from the lowest income level as compared to other program sites, immersion strategy only, early-exit only, and late-exit only program sites have roughly the same proportion of their student population from the two lowest income levels (i.e., low and low-middle) (87.9%, 83.4% and 83.8%, respectively). Proportionately more of the student body from program sites having both an immersion strategy and an early-exit program were from homes that were estimated to be middle income level or higher.

Site administrator estimates of the occupational levels of their students' heads of household were somewhat consistent with their estimates of family income (see Table 132). Proportionately more students in immersion strategy only sites than other program sites were judged to come from homes where the head of household was unemployed. Nonetheless, almost the same proportion of students in immersion strategy only and early-exit only program sites were judged to come from homes where the head of household was either unemployed or unskilled labor (67.2% and 68.4%, respectively). In contrast to their family income estimates, site administrators judged that proportionately more students from late-exit sites come from homes where the head of household was either a professional or held a white collar job. Program sites that have both an immersion strategy and an early-exit program seem to be more similar to immersion strategy only and early-exit only sites in having proportionately more of the student body from families where the head of household was estimated to be either unemployed or an unskilled laborer.

The proportion of students at each program site receiving free or reduced-price lunches as well as aid to families with dependent children (AFDC) are additional indicators of the socioeconomic level of the student body (see Table 133). Once again, clearly proportionately more students from immersion strategy only sites than any other program sites were estimated to come from homes that were receiving AFDC or were participating in the free or reduced-price lunch program (50.9% and 76.0%, respectively). While early-exit only, late-exit only, and immersion strategy/early-exit program sites had roughly two-thirds of their student body participating in the lunch program, almost two-fifths of students at early-exit sites were estimated to come from homes receiving AFDC (38.7%), as opposed to one-quarter of the students in late-exit only or immersion strategy/early-exit program sites (25.8% and 22.6%, respectively).

ERIC

Table 131

Mean Estimated Percentages of Student Body at Income Level, by Program

	Type of Program			
Income Level	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
Low	71.6	63.3	36.9	51.8
Low-Middle	15.9	20.1	28.2	32.0
Middle	12.0	13.9	27.9	9.0
High-Middle	0.5	2.1	3.5	5.8
Affluent	0.1	0.8	3.5	1.4

Table 132

Mean Estimated Percentages of Occupational Levels of Student Body Heads of Household, by Program

	Type of Program			
Occupation	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
Unemployed	30.2	23.6	24.5	20.1
Unskilled labor	37.0	44.8	36.7	15.7
Blue collar	23.3	14.0	23.6	33.7
Clerical	7.0	5.8	8.2	15.1
White collar	1.5	5.6	4.0	9.9
Professional	1.0	6.1	3.0	5.6

Table 133

Mean Percentage of Student Body Receiving Aid, by Program

	Type of Program			
Type of Aid	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
AFDC	50.9	38.7	22.6	25.8
Lunch program	76.0	69.1	68.3	64.1

In sum, immersion strategy only program sites, followed by early-exit only sites, consistently were estimated by site administrators to have the highest proportion of their student body from the lowest socioeconomic levels as compared to other program sites.

<u>How are students identified as being language-minority and limited-English proficient? Are there differences across programs?</u>

The process for identifying LEP students is largely a two-step process. The first requires identifying those students in the school population that are language minority. The second step requires the use of screening procedures to determine whether the English language skills of the identified language minority students are sufficient to allow their full and successful participation in a mainstream English-only classroom. There are alternative procedures to effect each of these two steps. All of the program sites in this study report having implemented such an identification and diagnostic process.

Six different identification procedures are used by study schools. The majority of students are identified as language minority either through a school office referral (77.5%) and/or a home language survey of all students (65.9%) as part of the regular school admissions procedures (see Table 134). Other procedures include home language surveys given to language minority families (57.3%), teacher initiated referrals (40.4%), and parent requests (36.9%). Program differences are noted. Immersion strategy only and early-exit only sites appear to use more of these procedures than do immersion strategy/early-exit and late-exit sites.

Table 134

How Students Are Identified/Who Initiates Process

	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %	Overall %
Home language survey of all students	76.5	100.0	40.0	35.9	65.9
Home language survey of IM students	85.8	86.8	40.0	16.0	57.3
School office initiates	100.0	86.8	40.0	69.9	77.5
Teacher initiates	62.3	47.0	57.0	14.4	40.4
Parents initiate	43.6	74.6	13.0	4.9	36.9

Schools consider a number of factors in their determination of which language minority students are limited English proficient. Overall, a student's English speaking and English comprehension skills, as well as the results of the home language survey are the three factors considered by most programs (95.1%, 95.1% and 89.0%, respectively) (see Table 135). This information is supplemented by data on the student's reading proficiency, parents' request, speaking and comprehension proficiency in the child's primary language, student ethnicity, teacher judgment, and reading proficiency in the home language. These factors are used by approximately two-thirds or more of the school sites in various combinations with one another. Program comparisons reveal that late-exit programs do not tend to use as many of the various factors named above in their identification of LEP students as do immersion strategy only, immersion strategy/early-exit, or early-exit only sites. Reflecting their respective program models, immersion strategy only programs do not consider the students' proficiency in their primary language as do sites offering bilingual instruction.

ERIC

Once students are identified as needing bilingual instruction, the majority of the schools (76.6%), across programs, provide services automatically, unless parents request that students not receive such services. However, language services are not always provided to all eligible students. Restrictions on services occur at times because of inadequate funds (7.0%), restrictions are sometimes placed on the number of years students can receive bilingual services (20.6%), restriction of services may be made to certain grades (3.1%), or limitations may exist on space available for bilingual classrooms (3.1%). Most of these restrictions, except for funding limitations, are reported only by early-exit schools.

ERIC

Table 135

Percentage of Schools Reporting Factors Used to Identify English Language Proficiency, by Program

	Type of Program				
Factors	Immersion %	Early-Exit %	Immersion-Early-Exit %	Late-Exit %	Overall %
Proficiency in speaking English	100.0	100.0	100.0	85.6	95.1
Proficiency in understanding spoken English	100.0	100.0	100.0	85.6	95.1
Home language survey	100.0	100.0	100.0	62.4	89.0
Parents' request	100.0	74.6	73.1	64.1	76.7
Proficiency: speaking and understanding home language	27.9	87.8	100.0	74.2	71.9
Proficiency in reading in English	100.0	62.4	43.9	91.2	77.7
Student's ethnicity	86.5	42.7	86.1	75.2	68.8
Teacher judgment	100.0	49.8	43.1	58.2	62.7
Proficiency in reading home language	27.9	72.7	43.9	78.0	61.0
Proficiency in writing English	67.4	43.7	43.9	58.2	53.6
Proficiency in English math	8.8	44.7	30.0	37.3	32.3
Proficiency in Spanish math	8.8	38.6	30.0	37.3	30.3

How are LEP students reclassified as FEP, and where are they placed after reclassification? Are there differences across programs?

Ninety-three percent of the schools reported formal policies for deciding when students should be reclassified as fluent in English (the only exceptions were late-exit schools) and, as with the initial assessment process, most schools use a variety of methods for reclassification. Approximately two-thirds (68.1%) reported that students are assessed at least once a year, and one-quarter (27.7%) of the sites review students' progress two to three times a year. One-third of the schools indicated that reassessment can be done at the discretion of school staff at any point in time during the year. Most commonly, reclassification procedures consider the child's ability to speak English and comprehend spoken English (see Table 136).

Despite the wide-spread reliance on testing, no school uses only a single method for reclassification. Reclassification is a complex process taking into account a variety of factors. While schools in each program reported using one or more of the factors listed, program differences are evident. Interestingly, not as many late-exit sites use English reading proficiency (66.5%), staff judgement (32.1%), teachers' judgment (61.7%), classroom grades (27.2%), or parent recommendations (16.0%) as criteria for reclassification as did the other programs. This contrasts markedly with schools in the other programs, wherein almost all, if not all, schools use English reading proficiency as a reclassification criterion. The most noticeable program difference is the finding that where all immersion strategy only school sites reported using classroom grades in reclassification, this was not done to the same extent in the other programs. Further, parent request does not play as large a role in reclassification in immersion strategy only sites as it does in the other program sites (50.2% vs. 76.8% to 81.7%). Notably, with one exception, proficiency in the students' home language does not play a significant role in reclassification in any program (31.5%). The exception is in the early-exit only sites where they use information on students' proficiency in speaking and understanding their primary language in the reclassification process (52.7%).

Table 136

Percentage of Schools Reporting Factors Used to Reassess English Language Proficiency, by Program

Method	Type of Program Immersion %	 Early-Exit %	 Immersion/ Early-Exit %	 Late-Exit %	Overall %
English oral/aural test	100.0	100.0	100.0	90.0	96.5
English reading proficiency	100.0	90.7	100.0	66.5	85.1
Parent request	50.2	81.7	80.1	76.8	72.8
Teacher judgment	100.0	90.7	80.1	61.7	82.2
English achievement/ content	100.0	60.1	81.4	61.7	70.8
Staff judgment	91.2	77.5	61.5	32.1	64.0
Classroom grades	100.0	66.7	52.7	27.2	60.0
Parent committee recommendation	32.6	53.1	80.1	16.0	38.9
District staff judgment	77.7	48.8	80.1	27.2	51.6
Home language oral/ aural test	27.9	52.7	19.9	20.0	32.3
Home language reading proficiency	27.9	31.5	19.9	20.0	25.5
Home language achievement/content	16.5	12.2	0.0	11.1	11.4

More than half the schools surveyed (55.4%) reported having formal policies for placement of students after their reclassification as fluent in English. Of those who have such policies, most (79.2%) said that students are transferred to regular, English-only classrooms. In some schools (19.7%) students continue to work in the same program classes. These possibilities exist in schools with all program types.

<u>What is the composition of classrooms in which target students are taught? Are there differences across programs?</u>

As noted earlier, the availability of native English speakers provides important learning opportunities for second language learners of English. To this end, data were collected from each program describing the classroom composition in terms of how the students at the study sites are mixed with native speakers of English. At kindergarten through third grade, sites with both an immersion strategy and an early-exit program provide the most opportunities for language minority students to be placed with native English speakers, either in single language minority groups with native speakers of English or in groups of mixed language minority groups with native English speakers (100.0% and 70.0%, respectively) (see Table 137). In contrast, late-exit only, immersion strategy only, and early-exit only sites tend not to mix their language minority students with native English speakers as much (≤74.6%). However, in grades four through six, there is much more mixing of language minority students and native English speakers, either all day for instruction (≥59.9%), or with only some pull-out for special instruction (50.1% to 69.8%) (see Table 138).

In sum, language minority students in grades four through six have more opportunities to mix with native English speakers than do language minority speakers in grades kindergarten through third grade. From a language development point of view, this pattern is not ideal. All programs should provide for greater mixing across grade levels.

Table 137

Percentage of Schools Reporting Grouping of Language-Minority and Non-Language-Minority Students, Grades K-3, by Program

	Type of Program			
Grouping	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
All students are from the same language-minority background	82.4	93.9	57.0	100.0
Students are from various language-minority backgrounds	10.9	12.2	43.1	14.1
Students from a single minority language background are mixed with English-language-background students	50.2	74.6	100.0	43.0
Students from various language-minority backgrounds are mixed with English-language-background students	32.4	43.7	70.0	10.5

Table 138

Percentage of Schools Reporting Grouping of Language-Minority and Non-Language-Minority Students, Grades 4-6, by Program

	Type of Program			
Grouping	Immersion %	Early-Exit %	Immersion/ Early-Exit %	Late-Exit %
Students are together for a full day's instruction (disregarding instruction in gym, music, or art)	59.9	92.5	100.0	75.7
Students are brought together from various homerooms, and they are together for two or more periods of instruction	11.1	34.0	0.0	38.4
Students are brought together from various homerooms, and they are together for only one period of instruction	12.5	31.2	18.8	18.7
Students are together for most of the day's instruction, but some are pulled out for short periods of special instruction	50.1	54.5	59.4	69.8

<u>What is the instructional program for language-minority students? Are there differences across programs?</u>

Almost all schools (88.0% of respondents) reported that they provide their limited-English-proficient students with the regular mainstream curriculum. For their special language programs, the majority of schools use separate materials from those of the base program (80.7%), although many use native-language versions of the same materials being used in English-only classrooms (58.3%). The schools use a variety of materials, with the majority of the schools using the same instructional materials for both special language and mainstream classes for some subjects (usually mathematics) and separate materials for others (such as social studies or language arts).

ERIC

Finally, immersion strategy programs differ from more traditional bilingual education programs in the degree to which they provide students with training in their home language and culture. All of the late-exit programs and all of the early-exit programs surveyed reported offering such training, but only two-thirds (66.4%) of the immersion strategy programs do so. When this instruction is offered (regardless of program), it is almost always (98.4%) integrated into other subjects such as language arts or social studies, rather than being offered as a separate subject or class (13.1%).

Respondents also were asked about how instruction for language-minority students is coordinated with that of monolingual English speakers. Generally (81.2%), such coordination is provided through regular — usually monthly — meetings, standardized teaching plans for all classes, and the daily impromptu discussions that occur among teaching staff. The majority of schools (81.9%) reported offering a planned program where the curriculum for LEP students is integrated with that offered to English-proficient students; schools often use resource staff to help coordinate instruction (78.9%).

<u>Are instructional objectives defined and monitored for each program</u>?

Almost all schools reported using formal, written instruction plans (skills continua) in English language arts, math, science, and social studies (see Table 139). Additionally, more than half the schools reported using such plans in Spanish for language arts and math, although fewer schools use skills continua in Spanish for science or social studies. Differences among programs are small, reflecting program goals such as no Spanish continua among immersion strategy program schools. Overall, English skills appear to be slightly better delineated than Spanish ones, possibly indicating a weakness in instructional goals and objectives among early-exit and late-exit programs.

Schools having continua generally adhere to these plans, with all reporting that teachers are required to use them. In most schools, classroom activities are monitored regularly, usually by the principal (98.5%) or sometimes through staff meetings (63.3%) or classroom visits (70.0%). Schools with late-exit programs are much less likely to use staff meetings for monitoring (26.8% compared to over 70.0% of schools with other program types).

Table 139

Percentage of Schools Using Written Skills Continua, by Subject and Language

Subject	English %	Spanish %
Reading	100.0	77.4
Oral language	91.3	70.0
Written language skills	100.0	67.0
Math	100.0	60.9
Science	96.1	42.6
Social studies	96.1	46.1

In addition to plans for classroom activities, almost all schools (97.8%) prepare written goals and objectives for students' progress and regularly monitor student progress through school records (76.2%), formal assessments by teachers and other staff (76.6%), and direct observation of students' behavior (84.9%) (see Table 140). For the most part, responsibility for monitoring pupil progress is the principal's or that of the classroom teacher. Other staff who contribute to monitoring pupil progress include resource teachers, district staff, and other school staff such as vice principals.

On the whole, students' progress is reviewed frequently, with over three-quarters the schools (78.2%) reporting that they monitor progress at least monthly.

Once information about student progress is collected, it is regularly relayed to parents and school staff through a variety of methods. The most common method of apprising staff of student progress is staff meetings (80.6% of respondents), although more than half the schools responding (60.9%) also distribute written reports to their staff.

To inform parents of students' progress, schools most commonly rely on parent-teacher conferences (90.4% of respondents), bulletins and newsletters (89.3%), and special programs such as parent group meetings (67.7%). Other methods of relaying information to parents include school site council meetings, school newspaper, and PTA meetings, as well as other, more informal meetings between parents and staff. Late-exit schools appear to make less use of conferences and meetings than other schools.

Table [illegible]

Monitoring of Student Progress

How schools monitor student progress:	Percent of Schools
Direct observation	84.9
Formal reports from teachers and staff	76.6
Regular school records	76.2
Informal reports (e.g., meetings)	67.1
Who monitors student progress:	
Principals	96.5
Classroom teachers	66.3
Resource teachers	58.5
District office supervisors/administrators	59.0
Vice-principals	46.1
Frequency of monitoring:	
Continuously	15.2
Weekly	17.7
Bi-monthly	15.1
Monthly	30.2
Quarterly	13.3
Semi-annually	3.5
Annually	5.0

Overall, schools with immersion strategy and/or early-exit programs appear to be quite similar in their approaches to monitoring instructional objectives. Late-exit schools use fewer methods of monitoring, and have fewer people responsible for monitoring progress than other programs. Late-exit schools are also more likely to communicate progress to parents through bulletins and newsletters rather than through conferences or meetings.

What is the level of parent participation in each program?

Most schools reported that they actively seek and encourage parent participation in their programs. Parent-teacher associations were reported to be active at most schools (91.5%), and more early-exit programs (89.8%) have parent advisory committees than do schools in late-exit (55.5%), or immersion strategy programs (55.5% and 65.0%, respectively). Schools also reported providing a variety of services for parents, with the most common being teacher-parent conferences, PTA meetings, school orientation classes and workshops, and parent counseling (see Table 141). There are no systematic differences in the services offered by schools with different programs.

Table 141

Percentage of Schools Offering Services for Language-Minority Parents

Type of Service	Percent of Schools
Parent-teacher conferences	93.9
PTA meetings	92.3
School orientation classes/workshops	86.5
Parent counseling	71.9
Ethnic heritage festivals	75.3
Community outreach program	64.3
Cross-cultural awareness classes/workshops	55.6
Adult ESL classes	25.4
Daycare or pre-kindergarten services	28.4

Schools said they encourage parents to visit classrooms, with one-third (34.6%) reporting an open door policy under which parents could

ERIC

visit classrooms at any time without appointments. Over half (55%) said they encourage visits but prefer to have parents make appointments.

Schools also reported that parents participate in various activities offered by the schools, most commonly: serving on school-level advisory committees, volunteering as classroom aides, providing schools with information about LM-LEP students, improving parent-staff communication, providing political and moral support to the school program, and raising money for the programs (see Table 142). About half the schools said they allow parents to work as classroom aides or to help plan the school curriculum for LM-LEP students. There are no systematic differences among programs in this regard, and overall, programs appear to be quite similar in their accessibility and level of parent participation.

<u>What are the characteristics of the communities in which these programs are implemented?</u>

Districts included in the study are located in areas which vary greatly in population (see Table 143). Two districts are located in cities with over one million people, two are located in areas with over 100,000 people, three are in cities with over 50,000 people, and the remaining two are in areas with fewer than 50,000 people. The population in the surrounding communities was reported as predominantly white, non-Hispanic (45.5%), with Hispanics representing 39.7%, blacks 11.7%, Asians or Pacific Islanders 2.5%, and other groups, 0.6%. School officials reported that their programs for LM-LEP students generally are well received in the community and most said they believed the parents of non-language-minority students hold generally positive attitudes toward language-minority students (84.5% positive) and toward the services for those students (75.5% positive).

Table 142

Percentage of Schools Reporting Parent Participation

Types of Parent Participation	Percent of Schools
Serve on school-level advisory committees	90.0
Serve as classroom volunteers	75.1
Serve as information resources for the school about LM-LEP students	77.9
Participate in activities to improve communication and interpersonal relations between staff and parents	74.3
Provide political and moral support to school and programs for language-minority students	73.1
Serve on district or area-wide advisory committee	74.4
Participate in fund raising	72.2
Serve as paid instructional aide	55.7
Assist with curriculum planning	52.9
Serve as volunteers for LM-LEP activities after school hours	32.5
Provide extracurricular native-language activities using school facilities	35.3

ERIC

Table 143

Population of School District Cities From 1980 U.S. Census of Population

Dist.	Total	White Non-Hisp. %	Hispanic %	Black Non-Hisp. %	Asian/ Pacific %	Other %
A	66,281	28.0	71.5	0.2	0.2	0.1
B	36,407	47.1	46.2	3.4	2.4	0.9
C	92,742	47.7	30.5	18.6	2.7	0.4
D	1,625,781	46.5	35.7	16.6	0.8	0.4
E	2,230,936	48.8	17.6	30.9	1.9	0.8
F	14,178	36.0	63.3	0.4	0.1	0.2
G	120,762	65.7	16.9	4.3	12.0	1.1
H	106,201	53.6	26.7	17.5	1.5	0.7
I	55,593	33.4	63.9	0.9	1.4	0.4

Included in "White Non-Hispanic" are individuals of American Indian, Eskimo, and Aleut background.

"Asian/Pacific" is made up of: Japanese, Chinese, Filipino, Korean, Asian Indian, Vietnamese, Hawaiian, Guanamian, and Samoan.

Summary

Forty-six schools were included in the study. Fourteen have late-exit programs in the study, sixteen have immersion strategy programs in the study, thirteen have early-exit programs in the study, and five have both early-exit and immersion strategy programs in the study. Most language-minority students in these schools speak Spanish at home, and about two-thirds of them are classified as LEP. Spanish-speaking students constitute between 34.4% and 76.5% of the students at their schools, with early-exit only and immersion strategy/early-exit schools having the greatest percentages of these students, and immersion strategy program schools the lowest. Most IM-LEP students in these schools are U.S.-born.

Late-exit, immersion strategy only, and early-exit only officials estimate that approximately three-fourths of the IM-LEP students entering their schools in September would still be there in June. Site administrators estimated that most students in their schools come from families with low to low-middle incomes, with immersion strategy schools estimating the lowest incomes. About two-thirds of the heads of households in immersion strategy only, early-exit only, and immersion strategy/early-exit sites are estimated to be unemployed or in unskilled occupations. In immersion strategy only schools, more families are estimated to receive AFDC and to participate in the school lunch program than families in the other school types.

All schools have formal procedures for assessing the language proficiency of language-minority students. Most schools use more than one assessment method, usually testing and home language surveys. In most cases, students classified LEP automatically receive language services unless their parents request otherwise. Most schools also have formal procedures for reclassifying students as FEP, and most have regular schedules for reviewing student progress. Again, although test scores are important in the reclassification process, all schools use more than one method to reassess students. In most cases, students reclassified as FEP are transferred to regular, English-only classrooms.

Schools having both an immersion strategy and early-exit program provide the most heterogeneous mix of students, with all of these sites mixing monolingual English speakers and language-minority Spanish speaking children. Most early-exit programs offer this mixture, but only half the immersion strategy only and two-fifths of late-exit programs do so. However, in all programs, more mixing of students was done in grades four through six than in grades kindergarten through three. From a language development point of view, mixing should occur at all grade levels.

Most schools use special materials for IM-LEP students, but most also try to provide these students with all aspects of the mainstream curriculum. Instruction for IM-LEP and FEP students is coordinated, usually through regular meetings and standardized teaching plans. Almost all schools have plans for English instruction in language arts, math, science and social studies, and more than half have such plans for language arts and math in Spanish. However, the lack of continuua in Spanish among a number of early-exit and late-exit sites suggests a potential weakness in not having a well-articulated and coordinated instructional program when Spanish is used.

Student progress is monitored by principals and teachers, and relayed to staff through staff meetings and parents through conferences and bulletins.

Most schools have active PTAs and most programs have parent advisory committees. Schools also offer a variety of services to parents and in most cases encourage parents to visit their children's classrooms. Most schools report that their language programs are well received in the community.

Except for differences relating to program models, overall, the programs in the study appear to be comparable in terms of their schools' administrative practices. There are some indications that schools with only immersion strategy classes are different from other schools, since these schools have more of their student body that are from low-income families and receiving AFDC, and more mobile students.

ERIC

District Context

To evaluate the larger context within which programs in the study operate, Aguirre International/SRA Technologies staff interviewed school district administrators at each of the nine districts participating in the study. Three of the districts have late-exit programs, while the other six districts are considered immersion strategy/early-exit districts (including the district which has only an immersion strategy program and the district which has only an early-exit program). The following discussion summarizes those interviews, and evaluates results to determine whether differences among districts result in systematic differences among programs unrelated to programmatic differences.

What is the size and distribution of the language-minority and limited-English-proficient population in the school districts? Does this differ among programs?

All nine districts use formal definitions of LM-LEP students; 29% of the districts use defined subcategories of LM-LEP students.

In each district, the number of students classified as LEP decreases as the students progress through school, but the rate of decrease is not consistent according to program type. That is, according to the underlying theory, immersion strategy and early-exit programs should concentrate on teaching students English so that they can transfer as soon as possible to English-only instruction. Thus, a fairly rapid drop in the numbers of LEP students in the higher grades is expected to occur. In late-exit programs, on the other hand, where concentration is theoretically on developing English skills after — or at least concurrently with — developing primary language skills, more students are expected to be classified still as LEP at higher grade levels and for the proportion so classified to decrease more slowly than in early-exit or immersion strategy programs. Tables 144 and 145 illustrate that these patterns do not appear to occur consistently among the districts surveyed. In fact, the early-exit/immersion strategy districts tend to have higher percentages of LEP students that have not been re-classified in the upper elementary grades than do the late-exit districts.

Table 144

Distribution of LM-LEP Students at each Grade as a Percentage of the Total School Enrollment, by Program

	Districts:								
	Late-Exit			Early-Exit/Immersion					
Grade	D	E	G	A	B	C	F	H	I
K	30.1	20.3	30.3	72.8	45.8	26.9	48.6	M	58.7
1	25.1	17.5	27.6	66.4	30.1	31.0	46.7	M	54.5
2	18.3	17.2	25.1	65.5	33.1	26.8	30.7	M	35.9
3	10.9	16.5	20.9	21.8	25.3	28.1	25.6	M	28.0
4	7.3	13.3	17.4	28.4	21.6	24.9	16.2	M	22.2
5	6.4	9.9	18.0	14.7	17.1	24.2	15.8	M	21.9
6	5.4	10.1	13.6	4.4	15.2	31.2	11.8	M	15.8

M = Missing; information not available

Numerous factors complicate any analysis of such data, especially considerations such as changes in the districts' reclassification criteria and influxes of LM-LEP students into the districts at various grade levels. Just such an influx regularly occurs, for example, in first grade, simply because in many districts, kindergarten attendance is not required. Table 145 displays the number of LM-LEP children, as well as the total number of children enrolled, reported by each district for each grade level. In District A it can be seen that another influx appears to have occurred at grade 4, where 412 students were classified as LEP as opposed to the 305 in grade 3. Noticeable jumps in population size such as this could result from any number of factors ranging from changes in immigration to shifts in school boundaries.

ERIC

Table 145

Distribution of Total Students and LEP Students by Grade, District, and Program

	K	1	2	3	4	5	6	All Grades
Late-Exit Districts								
District D:								
LM-LEP	4,534	4,353	3,175	1,960	1,330	1,183	1,005	17,540
Total	15,070	17,344	17,357	17,984	18,195	18,547	18,527	123,024
% LM-LEP	30.1	25.1	18.3	10.9	7.3	6.4	5.4	14.3
District E:								
LM-LEP	432	501	474	425	337	225	235	2,629
Total	2,126	2,860	2,748	2,574	2,543	2,282	2,319	17,452
% LM-LEP	20.3	17.5	17.2	16.5	13.3	9.9	10.1	15.1
District G:								
LM-LEP	430	424	367	315	258	280	208	2,282
Total	1,420	1,535	1,460	1,507	1,487	1,558	1,531	10,498
% LM-LEP	30.3	27.6	25.1	20.9	17.4	18.0	13.6	21.7
Early-Exit/Immersion Districts								
District A:								
LM-LEP	760	994	949	305	412	208	62	3,690
Total	1,044	1,497	1,449	1,397	1,450	1,414	1,396	9,647
% LM-LEP	72.8	66.4	65.5	21.8	28.4	14.7	4.4	38.3
District B:								
LM-LEP	469	410	339	276	218	148	123	1,983
Total	1,024	1,362	1,024	1,093	1,008	868	809	7,188
% LM-LEP	45.8	30.1	33.1	25.3	21.6	17.1	15.2	27.6
District C:								
LM-LEP	570	649	533	527	456	414	419	3,568
Total	2,122	2,095	1,988	1,875	1,832	1,713	1,341	12,966
% LM-LEP	26.9	31.0	26.8	28.1	24.9	24.2	31.2	27.5
District F:								
LM-LEP	153	178	118	97	59	59	41	705
Total	315	381	384	379	364	373	348	2,544
% LM-LEP	48.6	46.7	30.7	25.6	16.2	15.8	11.8	27.7

ERIC

Table 145 (continued)

	K	1	2	3	4	5	6	All Grades
Early-Exit/Immersion Districts								
District H:								
IM-LEP	M	M	M	M	M	M	M	M
Total	M	M	M	M	M	M	M	M
% IM-LEP	M	M	M	M	M	M	M	M
District I								
IM-LEP	319	324	202	156	122	113	85	1,321
Total	543	595	563	558	550	517	537	3,863
% IM-LEP	58.7	54.5	35.9	28.0	22.2	21.9	15.8	34.2

M = Missing; information not available

What is the district policy with respect to the role of English and Spanish in the classroom? Is this different for different programs?

When asked to categorize their use of English and the students' primary language in their programs for IM-LEP students, district officials differed noticeably, generally along program lines. All district administrators of late-exit programs and 83% of early-exit administrators indicated that English and their students' primary language are used equally for instruction. About three-fourths of immersion strategy program administrators reported that the students' primary language is used for support of English instruction.

District administrators also were asked if there was a district policy concerning the use of students' primary language. Across programs, most (89%) district administrators indicated that they had such a policy, and of these, all stated that the students' primary language could be used as the language of instruction. Over three-fourths (80%) of district administrators also reported that their districts had a policy regarding the teaching of content subjects. Of those that had such a policy, less than one-tenth (9.3%) said that content area instruction was delayed until students were English proficient. Fifty-four percent of the administrators said that the policy is that students would be taught content areas in English concurrently with learning English.

Programs differed in their policies, some of which were inconsistent with their respective instructional models. Slightly less than one-fifth (16%) of the early-exit/immersion strategy administrators reported a limit on the length of time IM-LEP students could receive special services addressing their language development needs. That is, some district administrators reported that LM-LEP students would only receive language support services for a limited number of years. At the end of the specified time period, the student would be mainstreamed regardless of whether the student had reached district criteria for reclassification. In contrast, almost three-fourths (74%) of the district administrators of late-exit programs reported such time restrictions to services (three and six years). In both instances such time limits are inconsistent with the pedagogic rationale underlying each instructional model. When the time limit expires, policy overrides pedagogy in determining the instruction of IM-LEP students. However, policy needs to be implemented at the school level by principals and teachers. A comparison of responses of district administrators with those of school site administrators in late-exit districts reported earlier reveals that the two do not agree. Earlier, we found that fewer than one-fourth of school site administrators reported such time limits, and that students were reclassified per instructional criteria. These differences between district and school site administrators suggests the different levels of concern and responsibility between the two administrative levels. District administrators are more concerned with policy issues, whereas site administrators and staff are more concerned with the instruction for each child within the context of district policy. It would appear that school site administrators tend to emphasize pedagogy over policy.

What is district policy with respect to program goals? Does this differ by program?

All district administrators agreed that their programs include goals of familiarizing students with American society and culture and providing the skills needed to function effectively in English-only classrooms. Apart from these goals, however, administrators differed about the purpose of their programs for LM-LEP students. Responses were offered for the individual programs, rather than for each district administrator, in order to assess each administrator's understanding of the separate programs for which he or she was responsible.

The most notable difference in district administrators' attitudes toward the purpose of their programs is that all (100%) late-exit administrators agreed that their program goals are to maintain and improve students' skills in their primary language, while no administrators said that this was a goal of immersion strategy programs and only half (48.9%) said it was a goal of early-exit programs (see Table 146). Roughly one-third of the late-exit administrators disagreed as to whether the program was designed to enable students to succeed in English-only classrooms, a goal expressed by administrators for all the immersion strategy and early-exit programs. Here again, the role of the primary language in instruction seems to differentiate the programs. All late-exit administrators and two-thirds (65.4%) of the early-exit administrators said that a program goal was to prevent LM-LEP students from falling behind EP students in content areas. These program differences are consistent with their respective instructional models.

Table 146

Percentage of District Administrators Agreeing With Program Goals of Instructional Services

Goals	Immersion %	Early-Exit %	Late-Exit %
To bring the English proficiency of IM-LEP students to the level necessary to function effectively in an English-only classroom.	100.0	100.0	60.1
To maintain and improve the primary language proficiency of IM-LEP students.	0.0	48.9	100.0
To familiarize IM-LEP students with American society and culture.	100.0	100.0	100.0
To provide the skills (other than use of the English language) necessary to function effectively in classrooms in U.S. public schools (test-taking skills, expected classroom behaviors, etc.).	100.0	100.0	100.0
To provide IM-LEP students subject matter content (math, social studies, etc.) in their primary language(s) so that they do not fall behind English-proficient students in these subjects.	0.0	65.4	100.0

<u>What is district policy with respect to the integration of IM-LEP and English-proficient students? Does this differ by program?</u>

In general, IM-LEP students spend at least part of their school day in classes with English-proficient students. In most (85.8%) districts, all students share classes in certain non-academic subjects such as music and art. In over half (56.9%) of the districts students are together for other instructional periods as well. In over one-third (39.9%) of the late-exit programs, but none of the other programs, students are grouped by language proficiency for instruction in language arts, so that IM-LEP students are separated from EP students during this content area. Notwithstanding, all programs provide opportunities to integrate IM-LEP and EP students.

What district administrative support is provided to each program?

In every district, services for LM-LEP students are coordinated at the district level, with each district having at least one full-time staff member assigned to this task.

While districts reported using a variety of funding sources to support their programs for LM-LEP students, it is interesting that Title VII funding was reported by all late-exit programs but only by about one-third (29.7%) of the early-exit/immersion strategy programs. Traditionally, Title VII funds have been provided only to early-exit programs; districts offering late-exit programs were able to obtain such funds to serve children while they are classified as LEP, but at the time that this study was initiated, immersion strategy programs were unable to obtain Title VII funding at all. Four-fifths (83.5%) of early-exit/immersion strategy programs are supported by local funds, while only two-fifths (39.9%) of late-exit programs are funded locally.

As can be seen from Table 147, the majority of staff providing services for LM-LEP students in each program are full-time teachers, with aides and tutors representing the next largest group of staff.

Almost all (96%) districts reported requiring their teachers of LM-LEP students to have either state or district certification in the teaching of LM students. Nevertheless, the lack of credentialed teachers forced many districts to waive their requirements. Sixty percent of late-exit districts reported that over 25% of its teachers working with LM-LEP students had received waivers of professional certification.

Table 147

Percentage Distribution of Types of Staff Teaching LM-LEP Students, by Program

Types of Staff	Immersion Strategy %	Early-Exit %	Late-Exit %
Teachers offering LM-LEP students services related to students' limited-English-proficiency:			
Full-time	64.2	67.6	81.4
Part-time	0.0	0.0	1.1
Special education teachers of LM-LEP students:			
Full-time	0.0	1.8	1.5
Part-time	0.0	0.0	0.0
Paraprofessionals (aides and tutors) for LM-LEP students:			
Full-time	11.5	2.0	9.5
Part-time	15.3	24.6	5.5
Resource or instructional support staff (e.g., resource teachers, curriculum developers, counselors) for LM-LEP students:			
Full-time	0.5	0.7	2.9
Part-time	8.6	3.3	0.0

In general, professional requirements for teaching LM-LEP students include a current state teaching credential (80% of the districts) or a provisional certification if a credential has not yet been completed. Most districts also require a proficiency in a second language (95.5%); 66.8% require demonstrated proficiency in English. In 46% of the districts offering early-exit/immersion strategy programs, teachers of LM-LEP students are also required to attend in-service training in special services for such students. In all districts, policies regarding tenure, promotion, and salary are the same for all teachers; in 40% of the districts, however, teachers of LM-LEP students are subject to special requirements regarding in-service training and continuing education.

Summary

One contrast among districts is the percentage of students classified as LEP, with most early-exit/immersion strategy districts rather consistently classifying higher percentages of students as LEP than late-exit districts, even in the upper elementary grades. Examination of reclassification data by program reveals that immersion strategy and early-exit districts, contrary to district policy of mainstreaming their students within two or three years, tend to maintain their students in their programs for at least four to five years (the duration of this study). Consequently, the larger numbers of identified IM-LEP students in the higher grades may reflect, in part, this delay in mainstreaming.

With regard to languages used for instruction, all administrators indicated that district policy permits the use of the students' primary language for instruction. For early- and late-exit programs, administrators said that both English and the students' primary language are used, and for immersion strategy programs administrators said that the language of instruction is predominantly or only English. Very few administrators said that content area instruction is delayed until students have a good command of English.

For late-exit programs, administrators identified development of students' primary language skills as a goal, while no one identified this as a goal of immersion strategy programs, and only half of the administrators said it was a goal of an early-exit program. Enabling students to succeed in an English-only classroom was identified as a goal for all the early-exit and immersion strategy programs, but only two-thirds of the late-exit administrators agreed. Students in two-fifths of the late-exit programs are grouped by language proficiency for language instruction, while students in the remaining programs are grouped more heterogeneously.

ERIC

In all programs, the majority of staff providing services for LM-LEP students are full-time teachers, with aides and tutors representing the next largest group. Although all districts require teachers to be certified teachers of LEP students, in many cases districts have had to waive these requirements for teachers in immersion strategy and late-exit programs.

District administrators' opinions on pedagogical issues varied considerably, in most cases apparently reflecting their personal views rather than the theories underlying the programs they administered. Administrators identified different goals for the different programs they administered, however, with the role of the students' primary language usually differentiating program goals. No systematic district differences were found which might be reflected as differences among program types, with the exception of percentages of students classified as LEP.

ERIC

V. STUDENT CHARACTERISTICS

Introduction

This chapter provides information on the students participating in the immersion strategy, early-exit, and late-exit programs. It includes home and community background information which was gathered through parent interviews, student demographic characteristics, and information on the students provided by their teachers. As with other data collection efforts, the intent is to determine the similarity between the three programs with respect to student background characteristics which might affect program success.

Home Background

Parent interview questionnaires were administered to 1,750 parent couples, single parents or legal guardians of target students. These questionnaires were administered by community members who had been trained by Aguirre International/SRA Technolcgies staff. Interviews were conducted in the students' homes, schools, and in some cases, by telephone. In most cases, the interviews were conducted in Spanish. Questionnaires were checked for completion before they were returned to Aguirre/SRA. Further editing was performed at Aguirre/SRA and, when necessary, staff instructed interviewers to contact parents again to clarify answers or to obtain omitted information.

The parent questionnaire was designed to obtain information on target students' families and backgrounds, including parents' involvement in and support of students' educational programs, language use at home and in the community, and parental opinions regarding the school. These and other topics are discussed below.

What are the families of target students like? Do they differ by program?

Size. Families of target students have, on the average, 5.1 family members living in the household and .2 unrelated persons. In a third of the cases the unrelated household member is a child. The typical household appears to be made up of two adults, three children, and for about 12% of the households, one or more unrelated persons. Nine percent of fathers are absent from the home. These figures remain fairly constant across programs.

Parents, residency and ages. About 45% of the students' mothers were born in Mexico and 13% were born in the U.S. Seventeen percent were born in Puerto Rico. Fathers' places of births are very similar to mothers'. As of their child's entry into kindergarten, mothers have lived here an average of 11.9 years and fathers, 14.2 years.

Sixty percent of the mothers are from 25 to 34 years old and 28% are between 35 and 44. Forty-five percent of fathers are between 25 and 34, and forty percent are between 35 and 44. There are no large differences among programs.

Target students' residency. As of their entry into kindergarten, target students have lived in the U.S. an average of 4.5 years, or for most of them, all their lives. There are few large differences across programs. Seventy-nine percent of target students were born in this country, 7% were born in Mexico, and 6% in Puerto Rico.

Parents' incomes. Parents reported their total gross family income. Overall, about 72% of the families report incomes under $15,000 (see Table 148). Remembering that the average family size is 5.1 persons makes it clear that many families have relatively small incomes, and in fact almost 10% have gross incomes under $5,000.

Table 148

Percentage of Parents Reporting Total Gross Family Income, by Program

Income Category	Program Immersion %	Early-Exit %	Late-Exit %	Overall %
< $2,500	4.5	3.1	1.5	2.9
2,500-4,999	4.4	7.2	8.8	6.9
5,000-7,499	16.5	9.4	30.4	19.9
7,500-9,999	11.7	13.3	17.2	14.3
10,000-12,499	19.1	15.5	13.9	16.0
12,500-14,999	12.1	13.2	10.3	11.7
15,000-17,499	6.2	9.7	5.3	6.9
17,500-19,999	6.6	7.2	3.7	5.6
20,000-22,499	5.6	7.2	3.1	5.1
22,500-24,999	3.8	5.0	1.5	3.2
25,000-27,499	1.4	2.6	0.2	1.3
27,500-29,999	2.8	3.0	1.0	2.1
> 30,000	5.5	3.5	3.1	4.0
Weighted Number of Responses	392.1	353.4	485.6	1231.1

Over 25% of responding families receive Aid to Families with Dependent Children (AFDC), 34.7% receive food stamps, and 13.6% receive some other kind of assistance (see Table 149). When income and assistance are compared across programs, immersion strategy and early-exit program families appear to have similar incomes, while a higher percentage of late-exit program families fall into the income ranges below $7,500 a year (40.7% late-exit vs. 25.4%, immersion strategy and 19.7%, early-exit). Fewer than 9% of late-exit program families have an annual income greater than $20,000 while more than twice that percentage of immersion strategy and early-exit families have incomes above $20,000 (19.1% and 21.3%, respectively). More late-exit program families (over 30% more) receive AFDC and over 22% more receive food stamps. (Please note that the finding regarding the relative income level of parents of target students is different than the finding based on site administrators' estimates of all families in their school sites. As indicated, site administrators' estimates of the income level of families of all students enrolled in their schools suggested that the immersion strategy program had a larger proportion of low-income families while proportionately less late-exit program families were estimated to be in the lowest income category. In contrast, when collecting income data directly from parents of target students, the largest proportion of low-income families were in the late-exit program, with more immersion strategy and early-exit program families in the middle to higher income ranges.)

Table 149

Percentage of Parents Reporting They Receive Assistance, by Program

	Program			
Type of Assistance	Immersion %	Early-Exit %	Late-Exit %	Overall %
AFDC	16.8	12.7	47.2	25.6
Food Stamps	29.2	23.6	51.4	34.7
Other	15.7	12.6	12.4	13.6
Weighted Number of Responses	582.2	581.5	582.0	1744.5

Parents' education and occupations. Mothers and fathers of target students have completed an average of 7.7 years of schooling each. There is little variation among programs. Parents also were asked for their occupations. Responses were classified on a scale of the Socioeconomic Index for Occupations in the Detailed Classification of the Bureau of the Census, 1950 (Reiss, 1961). This scale was developed using information from a 1947 study of occupational prestige by the National Opinion Research Center, in which 88 occupations were rated by 2,920 respondents. Respondents weighted each occupation's standing on a five-point scale ranging from excellent standing to poor standing. Arbitrary weights were assigned to these responses and responses then were averaged to rank order the occupations according to their standing.

The 88 occupations in the scale are not direcdly comparable to the Bureau of the Census 1950 Alphabetical Index of Occupations and Industries. To create a more complete scale, the creators of the decile scale used summary statistics of the education requirements and income distribution of occupations (found to predict the NORC ratings satisfactorily) to construct a socioeconomic index including occupations on the NORC scale as well as occupations omitted from that scale. The decile scale is the distribution of this socioeconomic index in the male experienced civilian labor force in 1950. Thus occupations scored 10 include the 10% of this population with the highest ranking occupations.

Parents who were not employed for pay (housewives, retirees, etc.) were not coded. The scale used takes into account both the type of job and the industry in which the job occurs. Thus an airplane mechanic has a rating of 8, higher than an auto mechanic who has a rating of 5. To give a rough idea of the scale, laborers and private household services workers are at the low end of the scale, usually rated from 1 to 3. Foremen, clerks, salesworkers and skilled laborers are generally in the 4-7 range. In the 8-10 range are managers and professional and technical workers.

Relatively few mothers (29%) are employed for pay. Of those, over half fall in categories 1-4, in the lower portion of the scale, and only about 18% are in the upper range of managers and professional and technical workers. Seventy percent of fathers report that they are employed for pay. Of these, two-thirds are in the 1-5 range and one-third in the 6-10 range. Thirty-eight percent of employed fathers are in the lowest two categories where laborers and unskilled workers are classified. The largest group (50%) falls in the intermediate range of categories 4-7, where skilled workers, clerks, and foremen are classified. Only 11% are in the top managerial and professional categories. The proportion of immersion strategy program fathers holding positions in the higher categories is smaller than either late-exit or early-exit program fathers (5.1% vs. 16.4% and 12.8%, respectively), and a larger percentage of immersion strategy program fathers hold jobs in the lower two categories than early- and late-exit program fathers (45.5% vs. 33.0% and 34.5%, respectively). Other than those noted, there are few differences across programs.

Have target students had preschool experience? Does this differ by program?

Thirty-seven percent of target students have attended preschool and 63% have not. Of those who attended preschool, 53% attended preschools where both English and Spanish were used, 32% attended preschools where English was spoken, and 15% attended preschools where Spanish was spoken. A higher percentage of immersion strategy program students attended preschool than either bilingual program students, and of students in all programs who attended preschool, more immersion strategy students attended English-speaking preschools, while more early-exit children attended preschools where both English and Spanish were spoken (see Table 150).

Table 150

Percentage of Target Students with Preschool Experience, by Program

	Program		
	Immersion %	Early-Exit %	Late-Exit %
Attended preschool	44.4	37.0	29.4
English spoken	37.3	25.9	31.7
Spanish spoken	11.1	16.8	17.0
Both spoken	50.7	57.3	50.8
Weighted Number of Responses	582.4, 256.6	582.2, 215.6	583.1, 173.2

To what extent are target children exposed to English and/or Spanish in the home and community? Does this differ by program?

Home language use. Parents were asked a number of questions about the language(s) used by them and their children at home. Over two-thirds of (70%) the parents report that their child uses only Spanish in speaking to the parents, but only one-third (37.7%) report that the child uses only Spanish with siblings. Roughly four-fifths (78.9%) of

ERIC

the parents report that they use only Spanish at home with their families. There are slight differences across programs in using only or mostly Spanish at home (see Table 151). Children in all programs are least likely to use all or mostly Spanish among themselves. In each program, the majority of parents tend to use Spanish when speaking with one another or their children. While most of their children tend to use mostly Spanish when speaking with their parents, it is inferred that children tend to use English among themselves. The latter indirectly suggests student facility and home use of English.

Table 151

Percentage of Parents Reporting That Only Spanish or Mostly Spanish is Used, by Program

	Program			
Context	Immersion %	Early-Exit %	Late-Exit %	Overall %
Child to parents	69.6	71.3	69.0	70.0
Child to siblings	30.9	43.7	38.6	37.7
Parent to parent	85.5	87.1	87.2	86.6
Parent to children	77.4	77.6	81.7	78.9
Weighted Number of Responses	578.5	583.2	581.0	1742.7

Written materials. Parents also were asked about written materials in the home. In immersion strategy and early-exit programs, higher percentages of parents receive English reading materials than receive Spanish (see Table 152). Higher percentages of late-exit parents receive Spanish reading materials than receive English. Overall, with the exception of English language newspapers, only about one-fourth to one-third of the immersion strategy and early-exit parents receive each kind of reading material. In contrast, from two-fifths to over one-half of late-exit parents receive reading materials in English and Spanish. The latter suggests stronger support for reading among families of late-exit program students than those in immersion strategy or early-exit programs.

ERIC

Table 152

Percentage of Parents Reporting They Regularly Receive Reading Materials, by Program

Material	Program			
	Immersion %	Early-Exit %	Late-Exit %	Overall %
English language newspapers	42.8	48.0	42.9	44.6
English language magazines	35.8	37.7	42.3	38.6
Books in English	34.0	36.8	42.7	37.9
Spanish language newspapers	28.0	27.7	54.1	36.6
Spanish language magazines	27.2	31.6	49.2	36.0
Books in Spanish	25.0	27.4	48.7	33.7
Weighted Number of Responses	582.2	583.2	579.8	1745.2

Television and radio. Nearly all of the respondents (99.3%) report that they have a TV in their home. Those who have sets report that the set is on an average of 32 hours a week, of which 9 hours are devoted to programs in Spanish. Parents report that children watch an average of 2.7 hours of television in English a day, and 0.5 hours a day in Spanish. Ninety percent of the respondents report having a radio, record player or tape player in the home. They report that the radio (or record or tape player) plays music about 12.3 hours a week, about 8.8 hours of which are in Spanish. Parents say their child listens to radio programming in Spanish an average of 0.4 hours a day and to English programming 0.6 hours a day. In general, children spend about 84% of their TV time watching English programming. When they listen to the radio, record player, or tape player, they listen to English almost two-thirds of the time. There are only minor differences among programs. Thus, there is strong support for English language skills through the media in all three programs.

Playmates. Seventy-five percent of parents report their child plays with siblings and 60% say their child plays with neighborhood children. Thirty-four percent of parents say their child plays with children who speak only Spanish, 31% say their child plays with children who speak only English, and 81% say their child plays with children who speak both languages (parents made as many responses as applied). Of these children, 54% speak about equal amounts of Spanish and English, 23% speak mostly Spanish, and 21% speak mostly English.

Parents were asked to estimate the amounts of time their child spends playing with children in each language group. They estimate that, on the average, their child spends 62% of his or her playtime with children who use both languages, 22% with Spanish-speaking children, and 16% with English-speaking children. Differences among programs are small. Overall, there appears to be equal support for English and Spanish through play activity.

Neighborhood language use. Parents were asked what languages their neighbors use at home. Most families live in neighborhoods where Spanish is spoken at home at least half the time, and differences among programs are quite small (see Table 153). Thus, students come from communities where there appears to be equal use of English and Spanish.

Table 153

Percentage of Parents Reporting Most Neighbors Use Language At Home, by Program

	Program			
Language	Immersion %	Early-Exit %	Late-Exit %	Overall %
More than 75% Spanish	38.1	46.6	39.4	41.4
More than 75% English	21.6	19.3	24.3	21.7
Equal amounts	37.2	32.0	34.7	34.6
Don't know	3.1	2.2	1.6	2.3
Weighted Number of Responses	575.3	583.2	576.6	1735.1

ERIC

What do parents know of and think about the educational services their child is receiving? Does this differ by program?

Parents were asked if their child's school has special classes to help children learn. Forty-three percent say yes, 16% say no, and 41% do not know. There are some differences across programs. Those parents who say that special classes are offered were asked to name or describe the classes. Overall, 96% name special education, 68% name bilingual education, 42% name ESL classes, 9% name a migrant program, 12% name Chapter 1, and 8% name Headstart. Thirty percent more late-exit parents name bilingual education than immersion strategy or early-exit parents. Parents who say that special classes are offered were asked if their child attends special classes. Overall, 69% of parents who say special classes are offered also say their child attends special classes. Seventy-five percent say their child is in a bilingual education program. Other responses are special education (4%), ESL (31%), migrant program (3%), and Chapter 1 (5%). These parents identify the purpose(s) of their child's program as helping the child to speak English (87%), helping the child to read English (85%), helping the child to write English (81%), helping the child develop motor skills (40%), and helping the child learn math (45%).

Parents who say their child receives special services were asked if those services are satisfactory. Seventy-eight percent say the services are very satisfactory, and 14% say they are moderately satisfactory. Three percent find the services unsatisfactory.

Parents were asked about information aside from report cards that they receive from school. Overall, between half and three-quarters of parents report receiving the information listed. With the exceptions of information about achievement goals and contact persons at the school, a higher percentage of late-exit parents consistently report receiving information (see Table 154).

Table 154

Percentage of Parents Reporting Information Received from School, by Program

Information	Program Immersion %	Early-Exit %	Late-Exit %	Overall %
Achievement goals	64.1	55.6	60.0	59.9
Behavior standards	66.8	65.8	76.1	69.6
Contact person at school	73.5	64.4	73.2	70.3
School's achievement progress	66.1	61.8	72.7	66.8
School's behavior progress	65.2	63.2	75.0	67.8
Weighted Number of Responses	581.0	583.2	580.9	1745.1

Parents of target students appear to be quite satisfied with their children's schools. Forty-two percent rate their school as excellent and 49% as good. Eight percent say the school is fair and only about half a percent rate it as poor. More late-exit parents rate the school as excellent and fewer rate the school as good. Other than this, there are only slight differences across programs.

Parents also were asked how satisfied they are with the help their school offers to Spanish-speaking children. Ninety-seven percent of parents say that they are very or somewhat satisfied. There are some small differences across programs in the level of satisfaction parents express. Slightly more late-exit parents are very satisfied than are immersion strategy and early-exit parents, and a few more immersion strategy parents say they are not satisfied with their school's help to Spanish-speaking children (see Table 155).

ERIC

Table 155

Percentage of Parents Reporting Level of Satisfaction with School's Help to Spanish-Speaking Children, by Program

	Program			
Satisfaction	Immersion %	Early-Exit %	Late-Exit %	Overall %
Very satisfied	65.0	67.3	80.7	71.1
Somewhat satisfied	31.0	29.5	18.1	26.1
Not satisfied	4.0	3.3	1.3	2.8
Weighted Number of Responses	567.3	569.2	579.7	1716.2

<u>To what extent are parents involved in their child's educational program? Does this differ by program?</u>

Parent involvement in the support of their children's learning has been identified as an important predictor of student achievement (Ramirez et al., 1989). To this end, data were collected from parents regarding the type and frequency of their involvement.

<u>Planning</u>. Parents who say their children are in special programs were asked whether they participated in the decision to place their children in those programs. Eighty-two percent of these parents say yes (63% of immersion strategy parents, 78% of early-exit parents, and 93% of late-exit parents).

<u>School visits</u>. Parents report an average of 2.8 visits to school during the school year to talk with staff or observe their child's class (see Table 156). The difference in maximum visits most probably is due to variations in volunteer programs.

Table 156

Average Number of Visits to School During School Year, by Program

	Program			
	Immersion	Early-Exit	Late-Exit	Overall
Average visits	3.3	2.7	2.5	2.8
Minimum	0.0	0.0	0.0	0.0
Maximum	60.0	40.0	38.0	60.0
Weighted Number of Responses	535.4	558.0	552.8	1646.2

Advisory groups. Overall about 8% of parents report that one or both of them participate on a school committee or advisory group. There are only slight differences among programs.

Homework. Eighty-five percent of parents report that their child has homework, and of these 77% report that they help with this homework, either by actively assisting the child or by making sure the child does his assignment. Most parents who help with homework report that they do this in Spanish.

There are some differences between programs with fewer early-exit parents (75%) reporting their child has homework compared to immersion strategy parents (83%) or late-exit parents (98%). More late-exit parents tend to monitor their child's homework and slightly more tend to help the child complete homework (see Table 157). More immersion strategy parents help their child with homework using English (30.3%) than do parents in either early-exit (19.7%) or late-exit (9.9%) programs. The higher proportion of late-exit parents monitoring and ensuring that their children complete their homework might be facilitated by the provision of homework in Spanish and/or the encouragement of use of Spanish for instruction by school personnel.

ERIC

Table 157

Percentage of Parents Reporting They Help with Homework, by Program

	Program			
	Immersion %	Early-Exit %	Late-Exit %	Overall %
Monitor completion of homework	52.6	52.9	74.0	59.8
Help complete homework	15.6	17.7	19.3	17.5
Do not help	14.8	4.7	4.7	8.1
Help using English	30.3	19.7	9.9	18.9
Help using Spanish	59.8	70.3	76.9	69.9
Help in both languages	9.9	9.9	13.1	11.2
Weighted Number of Responses	582.2, 371.5	582.1, 370.0	580.2, 513.8	1744.5, 1255.3

<u>How much time do target students and their families spend reading? Does this differ by program?</u>

Parents report that they have at home an average of 22 books that are not schoolbooks, and that they have read an average of 2.5 books in the last three months. They report that their child spends an average of 3.1 hours a week on leisure reading. Immersion strategy program parents appear, on the average, to possess fewer books but to read more than parents in the other programs (see Table 158).

Table 158

Reading Habits in Families of Target Students, by Program

	Program			
	Immersion	Early-Exit	Late-Exit	Overall
Books in home	20.4	23.0	23.3	22.2
Books read last 3 mo.	3.1	2.4	2.0	2.5
Hours/week child reads	3.3	2.8	3.3	3.1
Range of Weighted Number of Responses	571-577	571-582	576-582	1719-1738

Sixty-three percent of parents report that someone in the household reads to their child in Spanish. About 24% say this happens once

ERIC

a week or less, 27% say it happens two or three times a week, and 13% say it happens almost every day. Fifty-four percent of parents say someone reads to their child in English. Sixteen percent say this happens once a week or less, 25% report two or three times a week, and 13% say it happens almost every day. The only large difference across programs is that more late-exit parents than other parents report that someone reads to their child in Spanish (74.1% vs. 51.8% and 64.3%) and that this occurs more frequently than in the other programs.

Because parents reading to their children is known to be related to the children's academic success, these data were analyzed to determine the percentage of parents reading aloud to their children, regardless of language. As shown in Table 159, a slightly higher percentage of early-exit parents than immersion strategy or late-exit parents read to their children. Overall, parents who do read are more likely to read in either Spanish or English than to read in both languages.

Table 159

Percentage of Parents Reporting Children Are Read To, by Program

	Program			
	Immersion %	Early Exit %	Late Exit %	Overall %
Do not read	21.4	15.9	20.1	19.2
Read in one language	47.8	47.9	36.7	44.1
Read in both languages	30.7	36.2	43.2	36.7
Weighted Number of Responses	574.9	583.2	583.1	1741.1

<u>What opinions do the parents of target students have about services to Hispanic children in the public schools? Does this differ by program?</u>

Parents were asked what instructional languages should be used with Hispanic children who do not speak English when they enter the public schools. Overall, the largest percentage of parents think children should be taught in both English and Spanish throughout their school careers to develop both languages (see Table 160). The next largest group think children should be taught mostly in English, with Spanish used to help them understand. There are differences among programs, with many more late-exit parents than immersion strategy or early-exit parents favoring equal use of English and Spanish. Half of the immersion strategy parents favor using mostly English with Spanish to clarify though only about 35% of early-exit parents feel similarly. Only 10% of immersion strategy parents and fewer other parents believe that students who speak only Spanish should be taught in English only. Only a slightly lower percentage believe that these students should be taught mostly in Spanish at first and gradually introduced to English until they are taught in English only. Fewer than 1% believe they should be taught in Spanish only.

Table 160

Percentage of Parents Selecting Language of Instruction, by Program

Language	Program: Immersion %	Early Exit %	Late Exit %	Overall %
English only	9.7	6.9	2.8	6.4
Mostly English, Spanish to clarify	50.0	34.8	8.0	30.9
Both English and Spanish throughout school career	35.4	49.7	84.5	56.6
Mostly Spanish at first, gradually increasing English to English only	4.5	8.2	4.1	5.6
Spanish only	0.4	0.4	0.6	0.5
Weighted Number of Responses	577.3	583.2	583.1	1743.6

Parents were asked the extent to which schools should offer certain services to LEP students (see Table 161). Across programs, the majority of parents felt that extra instruction in English (93.1%) and bilingual teachers (93%) should be provided by schools. Most also felt that the schools should teach about the Hispanic students' home culture and history (80.9%) and help them develop their Spanish language skills (75.1%). Approximately two-thirds of the parents felt that schools should teach content in Spanish (63.3%). Over half felt that teachers should be permitted to teach in Spanish (58.3%) and that Spanish should be allowed in the classroom (56.5%). Overall, only two-fifths (44.2%) felt that teachers should be required to teach only in English.

Table 161

Percentage of Parents Responding Extent to Which Schools Should Offer Services to LEP Children

Service	Great Extent %	Moderate Extent %	Limited Extent %	Not At All %
Extra instruction in English	47.4	45.6	6.3	0.7
Teach about home culture	32.0	48.9	16.7	2.3
Develop Spanish language skills	27.4	47.7	19.6	5.4
Teach content in Spanish	22.2	41.1	23.3	13.5
Bilingual teachers	62.1	31.0	5.1	1.9
Permit Spanish in classroom	19.9	36.6	26.8	16.7
Allow teachers to teach in Spanish	20.9	37.3	21.2	20.6
Require teachers to teach in English only	14.8	29.4	15.8	40.0

Range of Weighted Number of Responses 1595.3 - 1736.1

Differences across programs on how parents feel services should be provided tend to reflect, in some cases, the instructional model of each program (see Table 162). While the majority of late-exit (92.1%) and early-exit (72%) parents felt that schools should develop their children's Spanish language skills, less than two-thirds of immersion strategy parents agreed. However, the high proportion of immersion strategy parents reporting that the immersion strategy programs should help their children develop Spanish language skills suggests that the programs studied do not completely reflect the aspirations of language minority parents. Nonetheless, only slightly more than one-third of the immersion strategy parents felt that Spanish should be permitted in the classroom and that teachers be allowed to teach in Spanish, in contrast to over half of the early-exit and more than four-fifths of the late-exit parents. This apparent inconsistency between wanting to have schools help their children to develop Spanish language skills and yet wanting to limit the amount of instruction in Spanish may simply reflect parental concerns regarding their children's acquisition of English, or parents simply are expressing what they were told by school personnel would be in the best interests of their children. That is, while parents in all three programs want schools to help their children become proficient in English, they also would like resources allocated to the development of Spanish language skills.

Parents were asked whether they agree that children in bilingual programs in school are looked down upon. Nine percent agree and 91% disagree.

Table 162

Percentage of Parents Responding That Schools Should Offer Services to a Great or Moderate Extent, by Program

	Program			
Service	Immersion %	Early Exit %	Late Exit %	Overall %
Extra instruction in English	91.0	92.8	95.3	93.1
Teach about home culture	73.8	78.8	90.2	80.9
Develop Spanish language skills	60.9	72.0	92.1	75.1
Teach content in Spanish	47.5	57.6	84.6	63.3
Bilingual teachers	89.0	92.8	97.2	93.0
Permit Spanish in classroom	35.7	50.9	82.7	56.5
Allow teachers to teach in Spanish	37.2	51.2	86.1	58.3
Require teachers to teach in English only	59.2	45.7	29.5	44.2
Weighted Number of Responses	577.5	579.6	579.8	1736.1

Sixty-six percent of parents say they strongly support bilingual education programs, 30% say they support them, 2% oppose them, and 1% strongly oppose them. There are only small differences across programs.

Ninety-three percent of parents believe that children from Spanish-speaking homes should be encouraged to become part of both the English-speaking and Hispanic cultures. There are only small differences across programs.

Summary

The average target student lives in a family of two adults, and 3.1 children, with .2 unrelated persons living in the home. Most target students were born in the United States, although fewer than one-quarter of their parents were. Seventy-two percent of families of target students have gross annual incomes under $15,000, and 10%, under $5,000. Twenty-five percent receive AFDC, 35% receive USDA food stamps, and 14% receive other aid. More late-exit families than early-exit or immersion strategy families receive AFDC and food stamps.

The average parent has had almost eight years of education. Most mothers do not work for pay, and of those who do, over half have jobs at the bottom of the socio-economic scale. Most fathers are employed, usually in jobs in the low and middle socio-economic categories.

Thirty-seven percent of target students attended preschool. While overall about half attended schools in which both English and Spanish were spoken, proportionately more immersion strategy students than either early- or late-exit students attended preschools where only English was spoken.

Most communication between parents and children in the home is in Spanish, but with siblings, children use more English. Almost half the parents subscribe to an English language newspaper, and fewer than half receive a Spanish-language paper. Almost all families have television sets, and the children watch mostly English programming. When target children listen to the radio, tape player, or record player, about two-thirds is in English and one-third in Spanish. Most target children play with friends who use both Spanish and English, and most families live in neighborhoods where Spanish is spoken at least half the time. Thus, across programs there appears to be equal support for both English and Spanish.

Less than half of parents are aware of special programs or classes at their child's school, and rate their school as excellent or good. Almost all parents say they are very or somewhat satisfied with the help their school offers to Spanish-speaking children. More late-exit than other parents say they are very satisfied.

Most parents who say their children are in special programs say they participated in the placement decision. While few parents report visiting school during the year, or serving on a school committee or advisory group, they are involved with their children's learning. Most parents say their child has homework, and of these, over three-fourths report that they help with this homework. More late-exit parents than immersion strategy or early-exit parents help with or monitor their children's homework. As most parent assistance is provided in Spanish, it is suggested that the use of Spanish for instruction by the bilingual programs may somehow encourage these parents to become more involved with their children's learning, and result in greater satisfaction by late-exit parents with the instructional program provided to their children than among immersion strategy or early-exit parents.

Parents report an average of 22 books that are not school books at home, and say they have read an average of 2.5 books in the last three months. Eighty-one percent of parents report that they read to their children in Spanish and/or English.

Across all programs, most parents believe their children's schools should maintain both English and Spanish languages throughout their school careers. The next largest group of parents think teachers should use mostly English but use Spanish to clarify. These figures vary somewhat across programs. Parents believe their children should be provided with bilingual teachers and extra help in English. They also want their children to be taught their home culture, Spanish skills, to be taught in Spanish, and to have teachers allowed to teach in Spanish and who permit Spanish in the classroom. The majority of parents across programs want their children to know Spanish and English equally well. Thus, it appears that immersion strategy programs may not completely reflect the language goals of their students' parents.

Only 9% believe children in bilingual programs are looked down upon. Ninety-six percent of the parents support bilingual programs, and 93% believe Hispanic children should be encouraged to become part of both the Hispanic and the English-speaking cultures.

Student Data Base

Various descriptive data were collected on students in the study. These were compiled in the student data base. Some general descriptive characteristics of students who participated in the study and similarities or differences are discussed in this section.

What are the characteristics of the students in the study?

Seven hundred forty-nine (749) immersion strategy study students, nine hundred thirty-nine (939) early-exit students and six hundred sixty-four (664) late-exit students participated in the study. The largest cohorts from the programs were those which began the study as kindergarten students in fall 1984 (see Table 163). There were no entering third grade cohorts in either immersion strategy or early-exit programs, and there were no entering first grade cohorts in the late-exit program. Slightly more than half of the immersion strategy and early-exit students in the study were male and fewer than half were female, and slightly more than half of participating late-exit students were female. The average age of students at entry to the program in kindergarten was 5.3 years for children in the immersion strategy and late-exit programs and 5.4 years for students in the early-exit program.

Upon entry to the study, all kindergarten students in all three programs were classified as LEP (see Table 164). All immersion strategy first grade students were classified as LEP upon entry to the study, though some first grade early-exit and third grade late-exit students were classified as FEP when they entered the study (15.1% and 18.1%, respectively). Based on district reclassification criteria, some students were reclassified from LEP to FEP during the study. Unexpectedly, proportionately more students in the early-exit (12.6%) and late-exit (11.8%) programs were reclassified at the end of their kindergarten year than in the immersion strategy program (3.9%). Proportionately twice as many students in the early-exit (25.4%) and immersion strategy (21.2%) programs were reclassified at the end of first grade than in the late-exit program (12.7%) (see Table 165). Once again, proportionately more students in the early-exit and immersion strategy programs were reclassified at the end of second and third grades as compared to students in the late-exit program. However, contrary to expectations, proportionately more early-exit students were reclassified earlier than those in the immersion strategy program. Nonetheless, after 4 years in the program, approximately two-thirds of the immersion strategy and almost three-fourths of the early-exit students had been reclassified, versus about half of the late-exit students. It is not until after 6 years in the program that we find almost four-fifths of the students reclassified in the late-exit program. Surprisingly, and of great importance, only one-fourth of the immersion strategy students and less than one-fifth of the early-exit students were mainstreamed after 4 years in the program (see Table 166). This is noteworthy given that over two-thirds of the students in each program had been reclassified.

Table 163

Number of Students Participating in Study, by Program and Cohort

Cohort	Immersion		Program Early-Exit		Late-Exit	
	N	Weighted N	N	Weighted N	N	Weighted N
10: Fall, 1984, K	251	263	335	310	158	189
11: Fall, 1984, 1st	106	111	169	137	0	0
13: Fall, 1984, 3rd	0	0	0	0	101	121
20: Spring, 1985, K	19	18	113	40	94	105
21: Spring, 1985, 1st	76	62	57	20	0	0
23: Spring, 1985, 3rd	0	0	0	0	92	102
30: Fall, 1985, K	181	197	191	198	141	171
31: Fall, 1985, 1st	86	99	74	79	0	0
33: Fall, 1985, 3rd	0	0	0	0	78	95
40: Spring, 1986, K	11	13	0	0	0	0
41: Spring, 1986, 1st	19	22	0	0	0	0
Total	749	784	939	784	664	784

Table 164

Students' Language Status (LEP or FEP) at Entry to the Study, by Entry Grade and Program (Weighted)

			Program	
Entry Grade	Status	Immersion %	Early-Exit %	Late-Exit %
K	LEP	100.0	100.0	100.0
	FEP	0.0	0.0	0.0
1	LEP	100.0	84.9	*
	FEP	0.0	15.1	*
3	LEP	*	*	81.9
	FEP	*	*	18.1

Table 165

Percentage of Students Reclassified to FEP during the Study, by Years in Program

		Program	
Number of Years in Program	Immersion %	Early-Exit %	Late-Exit %
1 (End of K)	3.9	12.6	11.8
2 (End of 1st)	21.2	25.4	12.7
3 (End of 2nd)	37.9	43.8	28.0
4 (End of 3rd)	66.7	72.0	50.8
5 (End of 4th)	*	*	67.0
6 (End of 5th)	*	*	78.6

Note: Implications are slightly biased as LEP students who exited were dropped from the calculations while FEP students who exited were retained in the calculation. This tends to slightly increase the percentage of reclassified students as years in program increases.

ERIC

Table 166

Percentage of Students Mainstreamed during the Study, by Years in Program

Number of Years in Program	Immersion %	Program Early-Exit %	Late-Exit %
1 (End of K)	1.3	1.6	*
2 (End of 1st)	10.7	9.1	*
3 (End of 2nd)	19.4	14.0	*
4 (End of 3rd)	25.6	16.9	*

Note: Implications are slightly biased as LEP students who exited were dropped from the calculations while FEP students who exited were retained in the calculations. This tends to slightly increase the percentage of reclassified students as years in program increases.

Late-exit students had more absences per year than either immersion strategy or early-exit students (13.7, 11.9 and 9.8, respectively). Cohorts from the immersion strategy program participated in the study an average of 1.6 to 2.6 years; early-exit cohorts, 0.4 to 2.4 years; and late-exit cohorts, 1.9 to 2.7 years (see Table 167). The range in the average number of years in the program for early-exit program students extends lower than for immersion strategy and late-exit program students due to the loss of funding and subsequent discontinuance of the early-exit program for one cohort in one district in the study. The students in this cohort participated in the study for only one semester. More than half of the students from each program who were exited from the study were exited due to moving to a non-study school or district or for unknown reasons (see Table 168).

ERIC

Table 167

Mean Period of Study Participation in Years, by Cohort and Program (Weighted)

Cohort	Immersion	Program Early-Exit	Late-Exit
10: Fall, 1984, K	2.5	2.4	2.7
11: Fall, 1984, 1st	2.6	2.2	*
13: Fall, 1984, 3rd	*	*	2.5
20: Spring, 1985, K	1.6	0.4	2.6
21: Spring, 1985, 1st	1.6	2.3	*
23: Spring, 1985, 3rd	*	*	1.9
30: Fall, 1985, K	2.1	1.9	2.0
31: Fall, 1985, 1st	2.1	1.9	*
33: Fall, 1985, 3rd	*	*	2.2
40: Spring, 1986, K	2.2	*	*
41: Spring, 1986, 1st	2.1	*	*

ERIC

Table 168

Percentage of Students Exiting From Study, by Reason and Program

Reason	Immersion %	Program Early-Exit %	Late-Exit %
Transferred to Non-Study Class	25.0	27.7	22.8
Transferred to Non-Study School	8.0	8.0	7.5
Transferred to Non-Study District	43.7	35.5	44.6
Changed from Immersion to Early-Exit Program	3.7	*	*
Changed from Early-Exit to Immersion Program	*	1.3	*
Changed to Unknown Program	1.0	0.2	0.0
Mainstreamed without being reclassified to FEP	0.9	0.2	0.6
LE Mainstream	*	*	9.3
Placed in GATE program	0.5	0.7	0.0
Placed in Full-time Special Education Program	3.0	0.7	1.5
Absent more than 40 days within school year	1.1	0.9	0.9
Program ended	0.0	8.1	0.0
Unknown	13.1	16.9	12.9
Weighted Number of responses	423	486	414

Conclusions

Two thousand three hundred fifty-two (2,352) students in eleven cohorts participated in the study for an average of 0.4 to 2.7 years. Kindergarten students from all three programs were classified as LEP upon entry to the study, and the average age of participating students at entry to the program in kindergarten was 5.3 to 5.4 years of age.

While four-fifths of immersion strategy and early-exit students were reclassified after four years in the study, less than one-fourth of each group was mainstreamed. Evidently, while teachers and principals were reclassifying students, they did not feel that their students were ready for mainstream classrooms. This suggests that the exit criteria might be too low and that teachers judge that it takes more than four years in special support programs before LEP students have the skills to be mainstreamed.

Students And Their Instructional Program

One thousand six hundred forty-two student data sheets were obtained for students in the immersion strategy program, 1,847 for students in the early-exit program and 1,591 for students in the late-exit program (see Table 169). The student data sheets were filled out by each student's teacher after a site coordinator had oriented the teacher to the questionnaire. The sheets then were edited by the site coordinator and sent to Aguirre International/SRA Technologies, where further editing was performed. When necessary, teachers were contacted by Aguirre/SRA staff to clarify or complete responses.

ERIC

Table 169

Number of Target Students for Whom Student Data Sheets Were Collected, by Program and Grade

	Program							
	Immersion		Early-Exit		Late-Exit		Total	
Grade	N	Weighted N	N	Weighted N	N	Weighted N	N	Weighted N
K	413	430.2	577	521.0	352	378.4	1342	1329.7
1	607	643.0	612	585.1	247	265.6	1466	1493.7
2	390	398.7	373	348.5	215	230.4	978	977.6
3	192	185.5	216	184.0	386	403.3	794	772.8
4	40	35.6	69	53.6	187	201.2	296	290.4
5	*	*	*	*	136	145.8	136	145.8
6	*	*	*	*	68	68.6	68	68.6
Total	1642	1693.0	1847	1692.1	1591	1693.4	5080	5078.5

ERIC

The student data sheet was designed to elicit information on students' demographic characteristics, special needs and participation in special programs, their teachers' judgments of their abilities and difficulties, the amount of instructional time each student receives in English and Spanish, and the language proficiencies of each students' classroom peers. The student data sheet addresses six questions, which are discussed here.

<u>For whom were student data sheets collected and do they differ by program</u>?

Fifty-three percent of the students in the immersion strategy program for whom student data sheets were collected are male and 47% are female. In the early-exit program, half of the students for whom student data sheets were collected are male and half are female. Slightly more than half of the late-exit students for whom student data sheets were collected are female (53%).

Participation in a free or reduced price lunch program is used as an indicator of low socioeconomic status. Based on this indicator, collapsed across grades, the majority of student data sheets were collected for students who are low income. Nonetheless, there are marked program differences. Proportionately more late-exit students (91%) are of low income than students in the immersion strategy or early-exit programs (81% and 75%, respectively) (see Table 170).

Table 170

Percentage of Target Students Participating in Free or Reduced Price Lunch Program, by Program and Grade

	Program		
Grade	Immersion %	Early-Exit %	Late-Exit %
K	55.3	51.3	90.1
1	90.8	83.0	91.4
2	87.9	88.6	89.5
3	86.2	87.4	89.2
4	84.0	83.6	95.3
5	*	*	94.7
6	*	*	86.4
Weighted Number of Responses	1668.6	1673.6	1642.4

Do target students have special learning needs and, if so, what services are provided to meet them? Does this differ by program?

Table 171 shows the percentage of students in each program who are participating in other special programs as reported on their student data sheets. Roughly one-third of the immersion strategy and early-exit students and one-fifth to almost one-half of the late-exit students participated in Chapter 1 programs. Across grades, up to 23% of immersion strategy students participated in programs other than Chapter 1, as did 16% of early-exit students, and 11% of late-exit students. Thus, overall, a sizeable portion of target students across programs received services other than those provided through the study program.

ERIC

Table 171

Percentage of Target Students Enrolled in Special Programs, by Program and Grade

		Program		
Grade	Special Program	Immersion %	Early-Exit %	Late-Exit %
K	Chapter 1	33.3	32.0	28.1
	Migrant	2.3	2.4	0.0
	Special Education	0.0	0.5	0.3
	Other	20.3	14.1	0.7
1	Chapter 1	29.1	32.4	46.8
	Migrant	3.9	4.5	0.0
	Special Education	2.0	1.5	1.9
	Other	12.1	8.8	3.2
2	Chapter 1	32.7	41.1	44.3
	Migrant	5.2	5.0	0.0
	Special Education	5.6	3.2	0.0
	Other	10.6	8.0	8.9
3	Chapter 1	36.9	35.5	38.2
	Migrant	4.0	4.4	0.0
	Special Education	4.9	3.5	1.8
	Other	14.6	1.2	1.8
4	Chapter 1	31.0	36.9	28.0
	Migrant	5.2	3.2	0.0
	Special Education	7.6	9.7	4.4
	Other	0.0	0.0	1.5
5	Chapter 1	*	*	21.1
	Migrant	*	*	0.6
	Special Education	*	*	2.6
	Other	*	*	11.4
6	Chapter 1	*	*	34.9
	Migrant	*	*	0.0
	Special Education	*	*	6.8
	Other	*	*	5.8

Across programs, about the same number of students have special learning needs (see Table 172). Teachers indicated that more of these students tended to exhibit a speech impairment, learning disability or handicapped vision than other disabilities.

Table 172

Number of Children with Special Learning Needs, by Program and Grade

Grade	Need	Program: Immersion %	Early-Exit %	Late-Exit %
K	Deaf	0.0	0.8	0.0
	Mentally retarded	0.0	1.0	0.0
	Hard of hearing	0.8	3.8	2.4
	Speech impaired	9.2	15.7	12.6
	Visually handicapped	0.0	7.7	1.9
	Learning disabled	3.7	2.9	3.3
	Emotionally disturbed	0.0	0.8	0.0
	Orthopedically impaired	3.4	0.0	0.0
	Other health impaired	0.8	2.5	3.0
	Total:	17.9	35.2	23.2
1	Deaf	0.0	0.0	0.0
	Mentally retarded	0.0	0.0	0.0
	Hard of hearing	2.9	1.9	2.1
	Speech impaired	28.4	22.7	4.9
	Visually handicapped	11.6	4.3	4.0
	Learning disabled	23.0	13.5	4.9
	Emotionally disturbed	0.0	0.0	0.0
	Orthopedically impaired	0.8	0.8	0.0
	Other health impaired	1.4	0.0	1.2
	Total:	68.1	43.2	17.1
2	Deaf	0.0	0.0	0.0
	Mentally retarded	0.0	1.3	0.0
	Hard of hearing	1.4	2.5	3.3
	Speech impaired	14.0	10.1	4.5
	Visually handicapped	1.8	9.5	2.4
	Learning disabled	12.0	9.8	4.7
	Emotionally disturbed	0.8	2.9	1.2
	Orthopedically impaired	1.8	0.0	0.0
	Other health impaired	3.2	3.8	0.9
	Total:	35.0	39.9	17.0
3	Deaf	0.0	0.0	0.9
	Mentally retarded	0.0	0.0	0.0
	Hard of hearing	0.0	0.0	3.0
	Speech impaired	6.4	5.8	11.9
	Visually handicapped	1.8	7.4	11.3
	Learning disabled	10.7	5.4	11.0
	Emotionally disturbed	0.0	0.0	0.0
	Orthopedically impaired	0.0	0.8	0.0
	Other health impaired	0.0	1.5	8.0
	Total:	18.9	20.9	46.1

Table 172 (continued)

Grade	Need	Program		
		Immersion %	Early-Exit %	Late-Exit %
	Deaf	0.0	0.0	0.0
	Mentally retarded	0.0	0.0	0.0
	Hard of hearing	0.0	0.0	0.0
	Speech impaired	0.0	0.0	4.6
4	Visually handicapped	0.8	0.0	2.4
	Learning disabled	1.8	5.4	8.9
	Emotionally disturbed	0.0	0.0	0.0
	Orthopedically impaired	0.0	0.8	0.0
	Other health impaired	0.0	0.0	3.0
	Total:	2.6	6.2	18.9
	Deaf	*	*	0.0
	Mentally retarded	*	*	0.0
	Hard of hearing	*	*	0.9
	Speech impaired	*	*	2.8
5	Visually handicapped	*	*	2.1
	Learning disabled	*	*	4.0
	Emotionally disturbed	*	*	0.0
	Orthopedically impaired	*	*	0.0
	Other health impaired	*	*	3.3
	Total:	*	*	13.1
	Deaf	*	*	0.0
	Mentally retarded	*	*	0.0
	Hard of hearing	*	*	0.0
	Speech impaired	*	*	4.7
6	Visually handicapped	*	*	1.9
	Learning disabled	*	*	3.7
	Emotionally disturbed	*	*	0.0
	Orthopedically impaired	*	*	0.0
	Other health impaired	*	*	0.0
	Total:	*	*	10.3

Teachers identified students they considered to be hindered in their learning by various difficulties. Teachers in all programs see the highest percentages of students as hindered by lack of English proficiency, underdeveloped cognitive skills, and lack of parental involvement (see Table 173). First, second and third grade immersion strategy teachers see the highest percentage of students as being hindered by lack of English proficiency and underdeveloped cognitive skills. Across programs, a low percentage of students were identified as being hindered by a lack of motivation to learn or use English.

Do target students receive and complete homework? Does this differ by program?

Almost all target students in all classes except kindergarten are assigned homework (see Table 174). Across programs, thirty-five to fifty percent always complete their homework, 29% to 44% usually do, and approximately 8% to 18% complete it about half the time. Fewer than 15% rarely or never complete their homework. Overall, program differences are minimal.

Table 173

Percentage of Students with Impediments to Learning, by Program and Grade

Gr.	Problem	Program Immersion %	Early-Exit %	Late-Exit %
	Lack of English proficiency	30.2	33.1	4.1
	Lack of Spanish proficiency	6.7	12.3	5.5
	Lack of motivation to learn English	8.8	7.5	1.0
	Lack of motivation to use English	9.7	10.9	1.5
K	Lack of parent involvement	14.9	15.4	8.7
	Poverty background	14.4	15.6	7.0
	Underdeveloped cognitive skills	20.5	35.0	12.6
	Underdeveloped social skills	14.5	30.6	13.0
	Other	11.8	9.6	8.1
	Lack of English proficiency	34.0	14.0	20.2
	Lack of Spanish proficiency	4.6	6.8	3.9
	Lack of motivation to learn English	6.7	3.9	1.2
	Lack of motivation to use English	11.0	4.5	0.7
1	Lack of parent involvement	19.2	13.3	14.4
	Poverty background	21.8	7.6	13.7
	Underdeveloped cognitive skills	28.1	14.7	12.7
	Underdeveloped social skills	24.0	10.7	12.8
	Other	13.7	3.6	4.1
	Lack of English proficiency	28.9	30.0	14.6
	Lack of Spanish proficiency	4.5	8.3	7.1
	Lack of motivation to learn English	5.2	3.2	2.1
	Lack of motivation to use English	6.8	4.4	3.6
2	Lack of parent involvement	20.4	12.9	8.2
	Poverty background	18.4	11.4	4.1
	Underdeveloped cognitive skills	29.7	17.7	16.5
	Underdeveloped social skills	20.2	12.1	7.2
	Other	6.3	1.7	0.8
	Lack of English proficiency	39.3	35.7	33.0
	Lack of Spanish proficiency	3.4	5.5	9.5
	Lack of motivation to learn English	6.8	5.2	6.7
	Lack of motivation to use English	10.0	8.9	10.1
3	Lack of parent involvement	25.7	11.9	24.4
	Poverty background	18.1	11.8	26.0
	Underdeveloped cognitive skills	40.2	21.8	32.5
	Underdeveloped social skills	23.5	14.8	19.3
	Other	18.5	5.0	4.5

Table 173 (continued)

Gr.	Problem	Program Immersion %	Early-Exit %	Late-Exit %
	Lack of English proficiency	11.7	10.4	29.0
	Lack of Spanish proficiency	0.0	2.9	5.7
	Lack of motivation to learn English	2.4	2.0	7.0
	Lack of motivation to use English	2.4	10.5	9.8
4	Lack of parent involvement	9.0	11.4	25.0
	Poverty background	12.3	1[illegible].6	19.6
	Underdeveloped cognitive skills	17.5	11.2	24.2
	Underdeveloped social skills	10.2	7.9	16.7
	Other	0.0	7.7	0.6
	Lack of English proficiency	*	*	26.0
	Lack of Spanish proficiency	*	*	8.4
	Lack of motivation to learn English	*	*	8.4
	Lack of motivation to use English	*	*	5.1
5	Lack of parent involvement	*	*	23.5
	Poverty background	*	*	29.2
	Underdeveloped cognitive skills	*	*	23.8
	Underdeveloped social skills	*	*	13.0
	Other	*	*	7.3
	Lack of English proficiency	*	*	12.5
	Lack of Spanish proficiency	*	*	1.4
	Lack of motivation to learn English	*	*	1.4
	Lack of motivation to use English	*	*	1.4
6	Lack of parent involvement	*	*	7.7
	Poverty background	*	*	0.0
	Underdeveloped cognitive skills	*	*	5.5
	Underdeveloped social skills	*	*	4.1
	Other	*	*	13.0

Table 174

Percentage of Students Completing Homework Assignments, by Program and Grade

Grade		Program: Immersion %	Early-Exit %	Late-Exit %
K	None Assigned	24.9	34.4	5.7
	Always	25.0	20.8	49.8
	Usually	25.9	26.4	29.5
	About half the time	12.0	10.1	6.7
	Rarely	8.4	5.6	5.9
	Never	3.8	2.6	2.4
1	None Assigned	0.2	3.0	0.0
	Always	39.3	43.4	46.8
	Usually	34.9	33.9	33.2
	About half the time	15.1	9.2	10.2
	Rarely	8.2	8.7	8.5
	Never	2.3	1.8	1.4
2	None Assigned	0.0	0.2	0.0
	Always	40.2	43.1	46.7
	Usually	33.5	38.7	31.2
	About half the time	17.6	13.4	14.5
	Rarely	8.0	4.3	4.5
	Never	0.7	0.4	3.0
3	None Assigned	0.9	0.3	0.5
	Always	35.0	42.3	41.7
	Usually	41.9	44.3	34.7
	About half the time	11.1	8.1	14.2
	Rarely	9.1	4.3	8.4
	Never	2.0	0.7	0.5
4	None Assigned	0.0	0.0	0.0
	Always	50.2	41.9	36.9
	Usually	37.2	28.8	41.8
	About half the time	12.6	14.9	15.0
	Rarely	0.0	10.6	4.7
	Never	0.0	3.8	1.7
5	None Assigned	*	*	0.0
	Always	*	*	44.9
	Usually	*	*	37.5
	About half the time	*	*	12.5
	Rarely	*	*	5.2
	Never	*	*	0.0

Table 174 (continued)

Grade		Program		
		Immersion %	Early-Exit %	Late-Exit %
6	None Assigned	*	*	0.0
	Always	*	*	49.2
	Usually	*	*	38.5
	About half the time	*	*	11.0
	Rarely	*	*	1.4
	Never	*	*	0.0
Weighted Number of Responses		1688.8	1681.0	1679.1

What are target students' skills, and do these differ by program?

Teachers were asked to rate the skills of their students in English, Spanish, and mathematics. In cases where students had no skills and were not expected to have skills, teachers were instructed to rate the skills for those students "N/A".

Because proficiency ratings for students in the three programs are somewhat different, results will be presented separately. With few exceptions, the majority of immersion strategy kindergarten students are not expected to demonstrate proficiency in all of the English skill areas except vocabulary, pronunciation, reading mechanics and conversation (see Table 175). Teachers rate the largest percentage of students as "fair" or "good" in first through fourth grades in each area, and the proportions of students rated as "very good" are largest in grades two through four. With the exception of a small percentage of fourth graders who are not expected to be proficient in spelling, all fourth grade students are expected to and do show some proficiency in all areas.

ERIC

Table 175

Percentage of Immersion Strategy Students Rated at Proficiency Levels in English, by Grade

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Spelling	96.0	0.5	1.4	1.3	0.7	0.0
	Vocabulary	12.9	4.5	23.8	24.9	23.2	10.8
	Pronunciation	0.2	5.0	23.9	25.8	28.8	16.4
	Writing simple words	62.1	2.1	9.1	8.5	12.7	5.7
	Writing sentences	70.0	2.2	9.7	8.2	5.4	4.6
	Writing simple stories	75.4	4.7	7.8	5.7	3.7	2.8
	Reading comprehension	53.3	4.2	13.4	14.2	9.5	5.5
	Reading mechanics	8.2	10.1	26.3	19.9	23.4	12.2
	Conversation	0.2	4.1	25.5	27.1	25.7	17.5
1	Spelling	7.5	11.1	27.8	26.3	20.7	6.7
	Vocabulary	0.7	2.3	27.2	34.4	28.8	6.6
	Pronunciation	0.0	1.1	20.5	40.5	30.0	7.9
	Writing simple words	1.2	9.5	30.9	29.2	22.0	7.2
	Writing sentences	1.7	20.1	32.6	23.2	18.5	4.0
	Writing simple stories	14.7	23.5	23.9	24.3	10.7	2.9
	Reading comprehension	0.2	5.1	26.2	35.6	25.6	7.3
	Reading mechanics	0.2	3.4	29.2	32.7	26.0	8.5
	Conversation	0.0	1.2	16.6	37.0	34.9	10.4
2	Spelling	0.2	5.3	19.2	29.8	32.1	13.4
	Vocabulary	0.0	0.6	16.5	32.3	38.9	11.7
	Pronunciation	0.0	0.4	11.3	32.4	44.5	11.4
	Writing simple words	0.0	3.1	17.1	27.9	40.6	11.3
	Writing sentences	0.0	5.7	22.7	36.7	27.8	7.1
	Writing simple stories	1.0	11.8	27.9	35.4	19.8	4.2
	Reading comprehension	0.4	1.9	11.6	37.5	36.8	11.9
	Reading mechanics	0.0	1.8	15.8	35.8	36.5	10.2
	Conversation	0.0	1.0	8.7	28.1	46.0	16.2
3	Spelling	0.0	3.0	17.2	26.9	39.4	13.5
	Vocabulary	0.0	0.9	12.2	33.2	41.0	12.7
	Pronunciation	0.0	1.4	7.8	34.2	44.4	12.2
	Writing simple words	0.0	2.0	14.4	33.1	39.6	11.0
	Writing sentences	0.0	3.5	17.5	38.7	31.1	9.2
	Writing simple stories	0.0	5.6	25.1	41.4	20.1	7.8
	Reading comprehension	0.0	0.5	13.0	40.8	31.6	14.2
	Reading mechanics	0.0	1.9	9.8	36.1	39.2	12.9
	Conversation	0.0	0.9	6.7	24.8	53.1	14.5

Table 175 (continued)

Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
				Level			
4	Spelling	14.3	0.0	17.8	22.5	35.5	10.0
	Vocabulary	0.0	0.0	10.4	32.5	49.6	7.6
	Pronunciation	0.0	0.0	2.6	30.1	55.2	12.1
	Writing simple words	0.0	0.0	7.6	32.9	47.0	12.6
	Writing sentences	0.0	0.0	12.8	35.3	41.8	10.2
	Writing simple stories	0.0	0.0	17.8	37.2	34.9	10.2
	Reading comprehension	0.0	0.0	7.6	42.6	37.2	12.6
	Reading mechanics	0.0	0.0	7.6	38.1	42.0	12.3
	Conversation	0.0	0.0	5.0	17.5	62.6	14.9

The percentage of early-exit students who are not expected to be proficient remains fairly high in all grades except fourth for all English skill areas, with the exceptions of pronunciation and conversation (see Table 176). When the students of whom no proficiency is expected are removed from the group, the ratings tend to fall in the middle categories, though generally smaller proportions of early-exit students are rated as "very good" in English skills in grades kindergarten through third than are their immersion strategy counterparts; in fourth grade, higher proportions of early-exit than immersion strategy students are rated as "very good" across English skills.

Late-exit teachers do not expect a fairly high percentage of their kindergarten or first grade students to be proficient in English skills except pronunciation and conversation (see Table 177). Kindergarten, first, second and third grade late-exit teachers rate fewer of their students as "very good" in English skills than do their immersion strategy or early-exit counterparts. A range of seven percent to almost one-fourth of first and second grade students are rated as having no level of proficiency in English reading or writing skills. The proportion of students who are rated "very good" increases with grade from kindergarten to sixth, though the majority of third, fourth, fifth and sixth graders are rated as "fair" or "good".

Table 176

Percentage of Early-Exit Students Rated at Proficiency Levels in English, by Grade

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Spelling	98.0	0.6	1.1	0.3	0.0	0.0
	Vocabulary	1.8	7.9	46.4	27.9	13.3	2.8
	Pronunciation	4.9	8.4	39.9	30.7	13.9	2.2
	Writing simple words	77.5	6.0	11.1	2.8	2.7	0.0
	Writing sentences	84.5	8.9	4.8	1.1	0.7	0.0
	Writing simple stories	93.8	6.1	0.2	0.0	0.0	0.0
	Reading comprehension	78.6	3.8	10.2	4.7	2.0	0.7
	Reading mechanics	33.9	12.7	22.7	16.3	12.0	2.4
	Conversation	3.2	8.3	42.5	27.3	15.3	3.4
1	Spelling	44.2	7.0	16.8	14.1	11.1	6.8
	Vocabulary	11.1	1.8	20.7	36.0	24.1	6.3
	Pronunciation	9.5	2.1	20.6	35.9	25.8	6.1
	Writing simple words	37.9	5.5	15.2	20.4	15.6	5.3
	Writing sentences	41.9	6.3	20.6	17.3	10.2	3.7
	Writing simple stories	51.9	9.5	19.6	11.6	6.5	1.0
	Reading comprehension	39.1	4.2	11.6	21.5	18.0	5.6
	Reading mechanics	35.6	3.9	15.0	18.2	20.8	6.6
	Conversation	4.6	3.5	20.8	30.7	33.4	7.0
2	Spelling	30.9	4.8	11.4	22.0	24.4	6.5
	Vocabulary	15.6	1.0	17.0	29.6	31.8	5.1
	Pronunciation	7.8	1.1	14.1	36.7	35.8	4.6
	Writing simple words	24.5	4.7	13.4	22.5	29.4	5.5
	Writing sentences	25.7	6.6	16.3	27.1	20.8	3.4
	Writing simple stories	37.5	7.0	16.9	20.3	15.8	2.5
	Reading comprehension	29.7	5.3	13.2	25.5	19.8	6.5
	Reading mechanics	28.1	3.8	14.0	22.7	24.9	6.5
	Conversation	6.5	0.9	12.6	35.6	35.9	8.6
3	Spelling	18.9	3.3	22.4	22.6	20.5	12.4
	Vocabulary	14.6	0.3	14.7	32.8	28.8	8.9
	Pronunciation	2.7	0.3	17.5	33.7	35.1	10.8
	Writing simple words	12.2	5.4	25.5	20.2	29.4	7.4
	Writing sentences	12.7	9.0	24.3	23.8	22.8	7.5
	Writing simple stories	12.7	10.9	24.9	24.0	22.6	4.8
	Reading comprehension	14.3	7.1	19.0	24.5	23.5	11.6
	Reading mechanics	14.3	3.4	21.6	25.4	26.3	9.1
	Conversation	1.1	0.5	13.5	32.2	37.3	15.4

Table 176 (continued)

Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
				Level			
4	Spelling	1.8	0.0	25.8	19.5	29.4	23.5
	Vocabulary	1.8	0.0	8.2	35.2	40.1	14.7
	Pronunciation	0.0	0.0	9.4	33.3	35.9	21.4
	Writing simple words	1.8	0.0	13.1	28.6	33.8	22.7
	Writing sentences	1.8	0.0	15.1	25.2	43.7	14.3
	Writing simple stories	4.7	0.0	23.5	26.2	33.5	12.1
	Reading comprehension	1.8	0.0	14.1	33.2	32.8	18.1
	Reading mechanics	1.8	0.0	15.9	25.7	45.5	11.1
	Conversation	0.0	0.0	8.0	18.1	49.7	24.3

Table 177

Percentage of Late-Exit Students Rated at Proficiency Levels in English, by Grade

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Spelling	89.5	5.3	3.8	1.1	0.0	0.3
	Vocabulary	22.6	8.9	44.5	17.1	5.2	1.8
	Pronunciation	17.9	6.5	42.3	21.3	7.4	4.6
	Writing simple words	74.4	7.3	9.9	4.7	2.7	1.0
	Writing sentences	87.1	4.9	4.0	1.4	1.6	1.0
	Writing simple stories	95.1	4.3	0.3	0.3	0.0	0.0
	Reading comprehension	79.9	5.7	9.5	2.7	1.1	1.1
	Reading mechanics	70.0	6.2	11.2	6.1	3.6	3.1
	Conversation	18.2	9.6	42.5	16.9	7.5	5.4
1	Spelling	66.1	10.7	10.4	3.6	7.5	1.8
	Vocabulary	23.9	8.4	39.7	15.6	11.0	1.5
	Pronunciation	14.4	9.2	43.3	20.5	11.5	1.1
	Writing simple words	57.9	14.4	13.3	5.0	7.7	1.8
	Writing sentences	62.1	19.2	6.6	4.8	5.7	1.8
	Writing simple stories	63.1	22.3	5.2	4.1	3.9	1.4
	Reading comprehension	61.7	16.2	8.7	5.5	5.7	2.2
	Reading mechanics	45.2	15.0	18.4	11.0	8.6	1.8
	Conversation	9.0	12.0	42.0	20.4	15.1	1.4
2	Spelling	37.4	7.5	23.0	17.6	11.5	3.0
	Vocabulary	11.4	2.6	33.0	33.0	17.0	3.0
	Pronunciation	5.5	1.1	28.8	35.9	25.7	3.1
	Writing simple words	26.1	7.6	32.0	17.6	13.9	2.7
	Writing sentences	33.2	11.8	27.1	16.4	9.3	2.2
	Writing simple stories	33.9	23.9	24.8	7.1	8.9	1.4
	Reading comprehension	34.4	10.2	28.5	12.7	10.7	3.5
	Reading mechanics	21.5	7.4	37.9	20.4	9.4	3.5
	Conversation	2.5	1.1	28.3	33.9	28.4	5.8
3	Spelling	3.9	6.5	22.6	29.2	31.3	6.5
	Vocabulary	1.4	4.3	20.5	39.6	27.8	6.6
	Pronunciation	0.0	2.5	19.7	32.9	37.7	7.2
	Writing simple words	3.0	5.1	18.8	31.2	33.9	8.0
	Writing sentences	3.2	6.5	23.2	30.9	29.5	6.7
	Writing simple stories	3.9	8.4	28.1	27.8	26.5	5.4
	Reading comprehension	3.2	6.3	21.4	34.5	26.2	8.3
	Reading mechanics	3.2	4.5	22.4	29.5	32.2	8.1
	Conversation	0.0	1.4	19.3	32.5	38.2	8.6

ERIC

Table 177 (continued)

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
4	Spelling	0.0	1.2	18.5	41.6	28.1	10.6
	Vocabulary	0.0	0.0	15.1	45.0	28.8	11.1
	Pronunciation	0.0	0.0	11.7	42.7	34.5	11.1
	Writing simple words	0.0	1.8	19.1	33.5	33.8	11.8
	Writing sentences	0.0	2.5	23.3	37.7	25.2	11.4
	Writing simple stories	0.0	3.5	27.2	36.1	24.8	8.4
	Reading comprehension	0.0	1.2	16.4	42.0	28.5	11.9
	Reading mechanics	0.0	1.2	16.0	40.6	30.4	11.8
	Conversation	0.0	0.0	9.3	41.5	33.5	15.7
5	Spelling	3.2	0.8	12.0	37.9	34.3	11.8
	Vocabulary	3.2	0.8	10.1	38.3	35.3	12.4
	Pronunciation	3.2	0.8	5.4	36.0	40.3	14.2
	Writing simple words	3.2	0.8	9.0	28.9	40.9	17.1
	Writing sentences	3.3	0.8	13.6	34.8	31.6	16.0
	Writing simple stories	3.2	0.8	15.7	42.6	24.9	12.7
	Reading comprehension	3.2	0.8	5.9	43.3	36.3	10.5
	Reading mechanics	3.2	0.8	7.4	40.0	34.2	14.4
	Conversation	1.7	0.8	5.9	35.5	40.7	15.5
6	Spelling	0.0	0.0	6.8	30.3	47.4	15.5
	Vocabulary	0.0	0.0	2.7	28.9	55.9	12.5
	Pronunciation	0.0	0.0	2.7	23.5	52.8	21.0
	Writing simple words	0.0	0.0	6.8	11.9	57.9	23.4
	Writing sentences	0.0	0.0	6.8	17.3	55.6	20.3
	Writing simple stories	0.0	0.0	9.9	32.7	41.9	15.6
	Reading comprehension	0.0	0.0	8.5	20.4	46.7	24.4
	Reading mechanics	0.0	0.0	6.8	19.7	53.6	20.0
	Conversation	0.0	0.0	2.7	17.7	52.5	27.1

Teachers also were asked to rate students on their proficiency in Spanish skills. With a few minor exceptions, the majority of immersion strategy teachers did not rate students as expected to have proficiency in Spanish skills, only rating half or fewer in kindergarten and first grade to none in fourth grade (see Table 178). Of those students rated, most were rated as "fair", "good" or "very good", especially in vocabulary, pronunciation and conversation. Most first grade students who received ratings were rated as having no proficiency in reading or

ERIC

writing skills in Spanish. In second and third grades, as in lower grades, more students were rated in vocabulary, pronunciation and conversation, and these generally were ratings of "fair" or "good".

More early-exit teachers generally expect students to show proficiency in Spanish skills than immersion strategy teachers, though in kindergarten and first grade, they rate fewer students in reading and writing skills than other skills (see Table 179). Early-exit teachers rate the largest proportion of their kindergarten, first, second and third grade students as having a "good" level of proficiency in Spanish skills. When Spanish skills are expected of fourth grade early-exit students, they are rated as having a "good" proficiency level in vocabulary, pronunciation and conversation, but lower levels of proficiency in other skill areas in Spanish.

Late-exit teachers' ratings of student proficiency levels in Spanish are quite different from either immersion strategy or early-exit teachers. With the exception of a few skills at the kindergarten level, almost all late-exit teachers say that all Spanish language skills are applicable for their students (see Table 180). Across grades, the majority of students are rated as "fair" or "good" in Spanish skills. The largest proportion of students receiving ratings of "very good" are in sixth grade.

Across grades except in kindergarten, nearly all students in each program are expected to show some proficiency in computation and number concepts (see Tables 181, 182 and 183). In all programs, the percentage of students expected to demonstrate proficiency in verbal problem solving increases with grade. Generally, the proportion of students rated at "beginning" levels in math skills decreases after first grade and the majority are rated as "fair" or "good" across programs and grades. Of interest is that more late-exit kindergarten and first grade students are rated as "very good" in all math skills than either immersion strategy or early-exit kindergarten and first grade students.

Table 178

Percentage of Immersion Strategy Students Rated at Proficiency Levels in Spanish, by Grade

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Spelling	96.9	0.0	0.4	1.9	0.8	0.0
	Vocabulary	54.8	0.5	3.7	9.5	21.1	10.3
	Pronunciation	49.6	0.5	3.7	10.0	22.0	14.2
	Writing simple words	93.1	0.0	1.0	3.1	2.9	0.0
	Writing sentences	96.3	0.0	0.8	2.0	0.9	0.0
	Writing simple stories	96.8	0.0	0.4	2.0	0.9	0.0
	Reading comprehension	91.8	0.3	0.7	2.5	4.7	0.0
	Reading mechanics	87.3	2.0	1.5	4.9	3.2	1.1
	Conversation	46.1	0.7	4.1	9.8	23.6	15.6
1	Spelling	76.2	12.7	3.4	5.0	2.2	0.5
	Vocabulary	57.8	2.6	2.8	13.8	18.7	4.4
	Pronunciation	53.6	1.6	3.4	13.5	20.8	7.1
	Writing simple words	77.2	11.8	4.0	4.4	1.9	0.8
	Writing sentences	78.9	13.4	3.5	3.9	0.0	0.3
	Writing simple stories	79.1	13.7	3.5	3.5	0.0	0.3
	Reading comprehension	77.0	13.0	4.8	3.6	1.1	0.5
	Reading mechanics	75.1	11.8	5.2	4.9	2.3	0.7
	Conversation	42.0	2.8	3.4	16.5	23.3	11.9
2	Spelling	88.0	6.8	3.9	0.4	0.8	0.0
	Vocabulary	76.6	0.0	2.3	8.3	8.1	4.7
	Pronunciation	68.6	0.4	1.5	11.1	12.0	6.4
	Writing simple words	88.0	5.0	5.7	0.0	0.9	0.4
	Writing sentences	88.5	7.6	2.6	0.4	0.8	0.0
	Writing simple stories	88.5	8.1	3.0	0.0	0.4	0.0
	Reading comprehension	88.0	4.1	4.8	2.2	0.8	0.0
	Reading mechanics	88.0	2.0	5.3	3.9	0.9	0.0
	Conversation	56.5	0.5	2.4	12.5	16.0	12.1
3	Spelling	91.7	2.1	6.2	0.0	0.0	0.0
	Vocabulary	85.6	3.9	2.9	1.0	6.6	0.0
	Pronunciation	81.4	2.8	4.6	0.0	11.2	0.0
	Writing simple words	91.7	6.2	2.1	0.0	0.0	0.0
	Writing sentences	91.7	7.3	1.0	0.0	0.0	0.0
	Writing simple stories	91.7	7.3	1.0	0.0	0.0	0.0
	Reading comprehension	91.7	2.1	5.2	1.0	0.0	0.0
	Reading mechanics	91.7	3.1	4.2	1.0	0.0	0.0
	Conversation	70.1	0.0	3.7	5.3	16.9	4.0

ERIC

Table 178 (continued)

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
4	Spelling	100.0	0.0	0.0	0.0	0.0	0.0
	Vocabulary	100.0	0.0	0.0	0.0	0.0	0.0
	Pronunciation	100.0	0.0	0.0	0.0	0.0	0.0
	Writing simple words	100.0	0.0	0.0	0.0	0.0	0.0
	Writing sentences	100.0	0.0	0.0	0.0	0.0	0.0
	Writing simple stories	100.0	0.0	0.0	0.0	0.0	0.0
	Reading comprehension	100.0	0.0	0.0	0.0	0.0	0.0
	Reading mechanics	100.0	0.0	0.0	0.0	0.0	0.0
	Conversation	100.0	0.0	0.0	0.0	0.0	0.0

Table 179

Percentage of Early-Exit Students Rated at Proficiency Levels in Spanish, by Grade

Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Spelling	90.5	1.3	6.3	1.7	0.2	0.0
	Vocabulary	0.3	0.5	8.9	34.6	45.6	10.0
	Pronunciation	3.9	0.6	7.5	28.8	50.3	9.0
	Writing simple words	56.5	10.2	14.2	5.8	11.0	2.3
	Writing sentences	87.8	6.9	3.2	0.7	0.9	0.5
	Writing simple stories	89.9	8.1	1.6	0.0	0.4	0.0
	Reading comprehension	62.9	2.4	12.2	11.7	8.1	2.7
	Reading mechanics	8.2	3.1	26.6	31.0	24.7	6.5
	Conversation	0.4	1.0	7.7	32.5	45.5	13.0
1	Spelling	18.0	5.3	20.7	21.1	24.8	10.2
	Vocabulary	5.9	1.0	7.6	22.6	45.2	17.7
	Pronunciation	5.9	1.0	7.1	22.2	45.5	18.4
	Writing simple words	12.1	5.8	19.1	25.1	26.7	11.2
	Writing sentences	15.3	12.8	23.5	21.5	18.0	9.0
	Writing simple stories	27.8	18.0	18.1	16.6	14.4	5.0
	Reading comprehension	11.7	5.7	16.1	22.5	31.0	13.0
	Reading mechanics	10.6	2.8	19.6	23.3	30.0	13.7
	Conversation	5.8	1.2	6.5	20.3	46.3	19.9
2	Spelling	8.9	2.7	13.8	24.3	30.8	19.5
	Vocabulary	7.2	0.6	8.7	23.2	42.5	17.9
	Pronunciation	6.5	0.0	5.0	20.1	45.3	23.1
	Writing simple words	8.9	4.0	11.8	18.6	34.0	22.6
	Writing sentences	8.9	6.3	13.4	22.0	31.6	17.9
	Writing simple stories	12.0	7.9	15.6	23.9	26.9	13.7
	Reading comprehension	7.4	3.3	14.6	17.2	38.6	18.8
	Reading mechanics	8.9	2.4	10.8	21.0	37.7	19.3
	Conversation	6.6	1.1	4.5	17.5	46.0	24.3
3	Spelling	37.8	1.0	3.8	18.8	29.7	8.9
	Vocabulary	37.6	0.0	3.6	13.5	35.9	9.5
	Pronunciation	36.6	0.0	1.8	13.5	30.6	17.5
	Writing simple words	38.8	1.0	2.5	13.9	35.6	8.2
	Writing sentences	38.5	1.0	4.5	15.4	36.5	4.1
	Writing simple stories	38.5	1.0	6.8	17.9	32.4	3.4
	Reading comprehension	36.0	1.0	4.5	17.0	34.5	7.0
	Reading mechanics	37.8	1.0	6.2	13.0	32.4	9.6
	Conversation	31.9	0.0	3.0	14.2	34.4	16.6

Table 179 (continued)

Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
4	Spelling	91.4	1.9	0.0	3.4	0.0	3.4
	Vocabulary	78.4	1.6	0.0	7.4	12.6	0.0
	Pronunciation	71.1	0.0	0.0	1.5	27.5	0.0
	Writing simple words	89.7	1.8	1.8	3.3	0.0	3.3
	Writing sentences	89.7	1.8	1.8	3.3	3.3	0.0
	Writing simple stories	96.1	2.0	2.0	0.0	0.0	0.0
	Reading comprehension	92.8	1.9	1.9	0.0	3.4	0.0
	Reading mechanics	84.1	1.7	1.7	9.4	3.1	0.0
	Conversation	69.6	0.0	0.0	10.8	15.5	4.1

(Level spans columns N/A through Very Good.)

Table 180

Percentage of Late-Exit Students Rated at Proficiency Levels in Spanish, by Grade

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Spelling	79.4	6.5	5.8	3.5	3.1	1.6
	Vocabulary	5.2	1.2	6.3	23.3	47.2	16.9
	Pronunciation	1.6	0.6	6.6	17.6	49.1	24.5
	Writing simple words	40.8	1.2	15.5	19.7	19.9	2.9
	Writing sentences	65.0	2.4	8.3	10.8	12.1	1.5
	Writing simple stories	83.5	6.9	7.0	1.5	1.1	0.0
	Reading comprehension	52.9	2.2	17.6	13.8	11.4	2.1
	Reading mechanics	12.5	4.2	27.5	19.8	22.6	13.5
	Conversation	0.0	1.7	3.0	15.3	46.9	33.2
1	Spelling	2.1	11.7	24.4	22.1	24.5	14.3
	Vocabulary	0.4	1.9	15.9	28.8	33.7	19.4
	Pronunciation	0.0	1.4	10.7	29.5	37.6	20.7
	Writing simple words	0.7	9.7	18.5	23.4	30.6	17.1
	Writing sentences	5.3	11.2	22.6	22.5	24.7	13.7
	Writing simple stories	18.4	20.6	14.5	17.2	21.5	7.7
	Reading comprehension	2.1	7.2	18.0	26.4	29.3	17.0
	Reading mechanics	1.6	5.9	16.9	24.2	33.6	17.9
	Conversation	0.4	1.2	9.8	22.0	45.2	21.5
2	Spelling	1.2	5.6	15.7	26.6	33.8	17.1
	Vocabulary	0.8	1.2	13.0	28.7	36.4	19.9
	Pronunciation	0.8	0.8	7.0	22.7	47.5	21.3
	Writing simple words	2.0	4.4	14.5	22.2	39.0	17.9
	Writing sentences	4.1	7.4	14.1	27.7	34.5	12.3
	Writing simple stories	7.4	9.7	18.2	30.8	28.0	5.8
	Reading comprehension	6.2	3.3	13.7	24.2	35.0	17.7
	Reading mechanics	1.2	4.6	17.2	25.0	32.7	19.3
	Conversation	0.8	0.5	7.2	19.8	45.3	26.3
3	Spelling	2.4	2.3	8.7	25.2	44.7	16.8
	Vocabulary	1.9	1.2	7.4	26.7	46.3	16.6
	Pronunciation	1.9	0.3	4.1	19.1	55.0	19.6
	Writing simple words	2.6	1.7	5.7	21.1	50.5	18.4
	Writing sentences	2.6	2.2	7.9	24.9	47.0	15.4
	Writing simple stories	2.6	3.6	9.7	31.8	41.6	10.7
	Reading comprehension	2.6	2.0	8.5	30.2	40.2	16.6
	Reading mechanics	2.6	0.8	8.2	25.4	45.7	17.3
	Conversation	1.6	0.0	4.1	20.2	53.4	20.6

Table 180 (continued)

Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
		Level					
4	Spelling	2.0	1.3	9.4	33.0	39.9	14.4
	Vocabulary	1.0	0.0	7.0	38.5	40.2	13.3
	Pronunciation	1.0	0.0	6.5	33.1	43.9	15.5
	Writing simple words	1.0	1.9	13.2	30.2	40.9	12.9
	Writing sentences	1.0	3.2	12.9	33.2	36.9	12.9
	Writing simple stories	1.0	3.2	16.1	37.7	31.0	11.1
	Reading comprehension	1.0	1.3	9.6	33.5	40.5	14.0
	Reading mechanics	1.0	1.3	9.2	29.5	43.9	15.3
	Conversation	1.0	0.0	5.1	28.2	48.4	17.4
5	Spelling	6.0	0.0	4.2	38.1	43.5	8.2
	Vocabulary	5.9	0.0	3.5	31.8	47.2	11.6
	Pronunciation	4.0	0.0	2.6	23.4	53.6	16.4
	Writing simple words	5.4	0.0	4.4	25.1	48.0	17.3
	Writing sentences	6.0	0.0	5.2	31.8	43.2	13.8
	Writing simple stories	6.0	0.0	6.9	39.9	38.5	8.8
	Reading comprehension	4.0	0.0	3.5	30.6	48.4	13.4
	Reading mechanics	6.0	0.0	4.3	27.7	45.5	16.5
	Conversation	3.9	0.0	2.6	28.7	49.2	15.7
6	Spelling	0.0	1.4	2.8	32.4	41.2	22.2
	Vocabulary	0.0	0.0	4.1	26.2	45.7	24.1
	Pronunciation	0.0	0.0	4.1	22.2	45.3	28.5
	Writing simple words	0.0	0.0	2.8	23.4	44.6	29.2
	Writing sentences	0.0	0.0	5.6	22.4	41.4	30.6
	Writing simple stories	0.0	0.0	5.6	32.2	39.3	22.9
	Reading comprehension	0.0	0.0	1.4	26.6	44.2	27.9
	Reading mechanics	0.0	0.0	2.8	23.8	48.7	24.7
	Conversation	0.0	0.0	4.1	18.0	50.6	27.2

Table 181

Percentage of Immersion Strategy Students Rated at Proficiency Levels in Mathematics, by Grade

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Computation	40.2	6.3	17.0	15.3	13.9	7.4
	Concepts of numbers and computation	0.3	3.3	25.7	23.9	31.2	15.6
	Verbal problem solving	28.9	10.5	17.8	21.6	12.7	8.6
	Concepts of geometry and proportion	12.1	5.6	19.2	20.2	30.7	12.2
1	Computation	0.0	1.7	16.8	30.4	37.9	13.2
	Concepts of numbers and computation	0.0	1.6	16.2	29.4	40.2	12.7
	Verbal problem solving	0.1	11.8	22.4	29.8	27.7	8.1
	Concepts of geometry and proportion	1.0	3.7	22.2	34.0	32.5	6.7
2	Computation	0.0	0.4	7.5	23.0	51.1	17.9
	Concepts of numbers and computation	0.0	0.6	8.0	21.0	54.4	15.9
	Verbal problem solving	0.0	2.3	13.2	34.7	40.3	9.5
	Concepts of geometry and proportion	0.0	4.0	11.9	33.7	42.3	8.2
3	Computation	0.0	0.0	6.8	28.8	52.1	12.3
	Concepts of numbers and computation	0.0	0.0	7.3	26.6	54.9	11.3
	Verbal problem solving	0.0	3.0	15.4	40.2	34.5	7.0
	Concepts of geometry and proportion	2.8	1.1	11.3	29.6	45.6	9.5
4	Computation	0.0	0.0	5.2	25.1	52.6	17.1
	Concepts of numbers and computation	0.0	0.0	2.6	35.3	50.0	12.1
	Verbal problem solving	0.0	0.0	12.8	44.6	37.7	5.0
	Concepts of geometry and proportion	0.0	0.0	18.2	27.3	54.6	0.0

Table 182

Percentage of Early-Exit Students Rated at Proficiency Levels in Mathematics, by Grade

		Level					
Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Computation	37.1	2.4	14.0	23.4	19.4	3.7
	Concepts of numbers and computation	0.0	2.2	22.2	33.4	36.4	5.7
	Verbal problem solving	14.4	7.4	24.8	30.1	19.4	3.9
	Concepts of geometry and proportion	12.3	5.9	23.5	27.6	26.3	4.4
1	Computation	0.0	1.7	12.5	23.3	47.6	14.9
	Concepts of numbers and computation	0.0	1.3	11.2	23.3	48.4	15.8
	Verbal problem solving	6.8	6.7	18.1	23.2	36.7	8.5
	Concepts of geometry and proportion	23.7	3.8	12.3	16.1	34.7	9.4
2	Computation	0.3	0.5	4.8	17.2	57.4	19.8
	Concepts of numbers and computation	0.3	0.5	4.9	18.1	58.7	17.4
	Verbal problem solving	2.2	1.6	9.3	25.4	51.7	9.9
	Concepts of geometry and proportion	5.5	1.3	8.8	20.9	55.9	7.6
3	Computation	1.0	0.0	3.2	22.9	51.0	22.0
	Concepts of numbers and computation	1.0	0.0	3.8	23.5	51.8	20.0
	Verbal problem solving	1.4	1.9	11.7	39.1	37.1	8.8
	Concepts of geometry and proportion	14.5	1.5	7.3	27.1	35.2	14.4
4	Computation	0.0	0.0	4.6	33.2	50.5	11.7
	Concepts of numbers and computation	0.0	0.0	7.4	31.2	50.1	11.3
	Verbal problem solving	0.0	0.0	19.7	42.5	28.8	9.1
	Concepts of geometry and proportion	58.0	0.0	6.5	19.9	15.6	0.0

Table 183

Percentage of Late-Exit Students Rated at Proficiency Levels in Mathematics, by Grade

Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
K	Computation	7.5	3.0	22.2	15.1	34.5	17.7
	Concepts of numbers and computation	1.9	2.2	25.8	17.8	33.3	19.0
	Verbal problem solving	14.2	3.8	22.9	21.9	22.7	14.6
	Concepts of geometry and proportion	19.4	13.0	10.4	15.3	27.4	14.6
1	Computation	0.4	3.6	13.2	24.0	40.6	18.3
	Concepts of numbers and computation	0.4	3.1	14.0	23.0	42.3	17.3
	Verbal problem solving	1.3	8.5	22.8	24.2	29.5	13.8
	Concepts of geometry and proportion	0.8	7.6	24.3	26.0	29.5	11.7
2	Computation	0.8	5.4	8.1	28.4	44.3	13.0
	Concepts of numbers and computation	0.8	1.3	12.6	28.0	44.7	12.6
	Verbal problem solving	0.8	2.2	16.7	31.4	39.8	9.1
	Concepts of geometry and proportion	3.3	0.4	17.4	29.9	43.1	6.0
3	Computation	0.7	0.3	5.9	28.6	51.2	13.4
	Concepts of numbers and computation	0.7	0.5	7.0	33.2	47.0	11.5
	Verbal problem solving	0.7	0.5	15.5	33.7	38.6	11.0
	Concepts of geometry and proportion	0.9	0.5	15.2	32.9	40.8	9.6
4	Computation	0.0	0.0	3.3	40.3	42.8	13.6
	Concepts of numbers and computation	0.0	0.0	3.3	41.7	43.1	11.9
	Verbal problem solving	0.0	1.0	20.8	41.6	26.4	10.2
	Concepts of geometry and proportion	6.2	0.5	17.1	39.8	27.8	8.7
5	Computation	0.8	0.0	2.7	37.7	45.2	13.5
	Concepts of numbers and computation	0.8	0.0	3.4	35.5	46.9	13.5
	Verbal problem solving	0.0	0.0	10.8	43.8	32.9	12.5
	Concepts of geometry and proportion	0.0	0.0	16.0	41.5	31.9	10.7

Table 183 (continued)

Gr.	Skill	N/A %	None %	Begin. %	Fair %	Good %	Very Good %
		Level					
6	Computation	0.0	0.0	4.1	22.2	51.8	21.9
	Concepts of numbers and computation	0.0	0.0	4.1	27.0	45.6	23.3
	Verbal problem solving	0.0	2.7	5.4	34.9	39.1	17.9
	Concepts of geometry and proportion	0.0	2.7	6.8	22.8	51.2	16.5

What is the typical instructional program of target students? Does this differ by program?

Teachers were asked to fill in for each student the amount of time per week the student receives instruction in various subject areas by the teacher, team teacher, aide, or resource teacher. The average percentage of instructional time received by students in the immersion strategy, early-exit and late-exit programs is reported in Table 184.

Immersion strategy students receive somewhat more instruction in English language arts than early-exit students receive, though the difference diminishes with grade. The proportion of time in English for immersion strategy students decreases from 63.6% in kindergarten to 52.5% in fourth grade and the proportion of time in English for early-exit students increases from 35.1% in kindergarten to 51.1% in fourth grade. With very minor exceptions, immersion strategy teachers do not allocate any instructional time to Spanish language arts. This is consistent with the program model. At every grade level, late-exit teachers allocate a smaller proportion of instructional time to English language arts than either immersion strategy or early-exit teachers. Though they devote less time to Spanish language arts in grades four, five and six than in lower grades, late-exit teachers spend proportionately more time on Spanish language arts instruction than do early-exit teachers.

Table 184

Average Percentage of Instructional Time for Each Academic Subject, by Program and Grade

Grade	Content Area	Program: Immersion %	Early-Exit %	Late-Exit %
K	English Language Arts	63.6	35.1	25.8
	Spanish Language Arts	0.0	29.9	39.9
	Social Studies	7.0	6.9	8.6
	Mathematics	23.2	22.1	18.4
	Science	6.2	6.1	7.3
1	English Language Arts	60.2	33.5	23.2
	Spanish Language Arts	2.2	24.2	41.7
	Social Studies	7.5	8.0	7.2
	Mathematics	23.5	26.7	20.6
	Science	6.6	7.6	7.4
2	English Language Arts	60.7	34.4	25.7
	Spanish Language Arts	0.1	25.4	28.8
	Social Studies	8.2	8.5	8.3
	Mathematics	24.1	23.1	18.7
	Science	6.9	8.7	8.5
3	English Language Arts	58.7	40.6	36.2
	Spanish Language Arts	0.0	18.2	28.9
	Social Studies	8.6	8.6	7.4
	Mathematics	24.5	24.6	19.8
	Science	8.2	8.1	7.8
4	English Language Arts	52.5	51.1	43.0
	Spanish Language Arts	0.0	0.9	18.0
	Social Studies	18.0	14.6	9.1
	Mathematics	17.5	20.9	19.5
	Science	12.0	12.5	10.4
5	English Language Arts	*	*	42.9
	Spanish Language Arts	*	*	18.2
	Social Studies	*	*	9.4
	Mathematics	*	*	20.2
	Science	*	*	9.4
6	English Language Arts	*	*	41.0
	Spanish Language Arts	*	*	18.1
	Social Studies	*	*	11.7
	Mathematics	*	*	18.3
	Science	*	*	10.9

Early-exit, late-exit and immersion strategy teachers reported that their students receive approximately equal instructional time in math, science and social studies.

Teachers recorded for each student whether or not the student receives ESL instruction. Across grades, roughly seventy-nine percent of immersion strategy and early-exit children and 81% of those in late-exit receive ESL instruction (see Table 185). In all three programs, the proportion of students receiving ESL instruction generally decreases with grade, and between programs, more late-exit students receive ESL instruction than students in the other programs.

Table 185

Percentage of Students Receiving ESL Instruction, by Program and Grade

	Program		
Grade	Immersion %	Early-Exit %	Late-Exit %
K	87.7	86.8	94.5
1	79.8	82.1	96.7
2	70.8	79.2	94.0
3	76.1	60.2	73.4
4	57.1	49.2	71.5
5	*	*	48.8
6	*	*	32.6
Weighted Number of Responses	1634.6	1615.1	1608.9

In what language do teachers instruct target students? Does this differ by program?

In all three programs, more students receive instruction in their primary language from the main or responding classroom teacher than from other classroom staff (team/resource teacher or aide) (see Table 186). More late-exit students receive instruction in their primary language than students in the other programs. The proportion of these students remains high across all grades, in contrast to a decrease in the immersion strategy and early-exit programs. Teachers also were asked to record the extent to which English or the student's primary language is used to teach each of several subject areas. Most immersion strategy students are taught in English with primary language to supplement and English only, as are early-exit students (except in kindergarten) (see Tables 187 and 188). Consistent with the program models, more early-exit students than immersion strategy students receive instruction in their primary language only, though, without exception, no fourth grade early-exit students receive instruction in their primary language only or primary language supplemented with English. Thirty-five percent to over 60% of late-exit students in kindergarten through second grade are taught math, science, social studies and cultural heritage in their primary language only (see Table 189). The percentage of late-exit students receiving instruction in equal amounts of primary language and English tends to increase with grade, and the percentage of students receiving instruction in primary language only generally decreases. Except in sixth grade, the proportion of students taught in their primary language with English to supplement remains substantial, particularly in math, science and social studies. These patterns tend to hold true over all subject areas.

Table 186

Percentage of Students with whom Primary Language is Used in Teaching, by Program and Grade

Grade	Need	Program: Immersion %	Early-Exit %	Late-Exit %
K	Responding teacher	57.8	80.8	96.5
	Team/Resource Teacher	1.2	42.6	18.5
	Aide	17.0	50.1	50.6
1	Responding teacher	38.9	81.9	94.6
	Team/Resource Teacher	1.2	19.1	26.3
	Aide	28.6	45.8	52.3
2	Responding teacher	22.0	67.1	90.0
	Team/Resource Teacher	2.5	16.6	18.4
	Aide	16.0	38.1	25.6
3	Responding teacher	27.8	55.2	87.2
	Team/Resource Teacher	7.0	11.1	22.5
	Aide	18.2	29.4	23.9
4	Responding teacher	19.1	62.9	91.2
	Team/Resource Teacher	21.4	1.0	25.0
	Aide	0.0	1.8	9.3
5	Responding teacher	*	*	94.8
	Team/Resource Teacher	*	*	12.2
	Aide	*	*	3.9
6	Responding teacher	*	*	78.0
	Team/Resource Teacher	*	*	23.8
	Aide	*	*	2.8
Weighted Number of Responses		1680.8	1671.4	1666.1

Table 187

Percentage of Immersion Strategy Program Students Receiving Instruction in Language(s), by Grade

Gr.	Languages	Math %	Science %	Subject Social Studies %	Ethnic Heritage %	Other %
K	Primary only	0.0	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.2	0.0	0.0	5.7	0.0
	Equal primary and English	0.0	3.4	3.4	16.0	0.0
	English w/primary to supplement	58.4	36.3	38.6	35.9	25.8
	English only	41.4	46.9	43.5	29.3	42.1
	Not taught	0.0	13.5	14.5	13.1	32.2
1	Primary only	0.9	0.9	0.9	0.0	0.0
	Primary w/English to supplement	0.5	0.4	0.6	4.1	0.0
	Equal primary and English	0.3	0.6	0.6	7.4	0.0
	English w/primary to supplement	60.4	50.3	57.7	57.4	23.8
	English only	37.9	39.8	40.2	18.9	17.0
	Not taught	0.0	0.0	0.0	12.3	59.2
2	Primary only	1.1	1.1	1.1	1.1	0.0
	Primary w/English to supplement	0.0	0.0	0.0	1.1	0.0
	Equal primary and English	0.0	2.9	2.9	6.3	0.0
	English w/primary to supplement	42.4	42.0	37.4	32.8	15.2
	English only	56.5	54.0	58.6	48.3	26.2
	Not taught	0.0	0.0	0.0	10.4	58.7
3	Primary only	0.0	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	0.0	0.0
	Equal primary and English	4.7	4.8	4.8	7.1	0.0
	English w/primary to supplement	35.0	33.6	29.4	13.1	6.2
	English only	60.4	60.1	60.1	45.7	31.0
	Not taught	0.0	1.6	5.7	34.1	62.8

Table 187 (continued)

Gr.	Languages	Math %	Science %	Subject: Social Studies %	Ethnic Heritage %	Other %
4	Primary only	0.0	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	0.0	0.0
	Equal primary and English	0.0	0.0	0.0	0.0	0.0
	English w/primary to supplement	0.0	40.5	19.1	0.0	0.0
	English only	100.0	59.5	81.0	59.5	49.1
	Not taught	0.0	0.0	0.0	40.5	50.9

ERIC

Table 188

Percentage of Early-Exit Program Students Receiving Instruction in Language(s), by Grade

Gr.	Languages	Math %	Science %	Subject Social Studies %	Ethnic Heritage %	Other %
K	Primary only	22.0	7.6	7.6	7.6	38.7
	Primary w/English to supplement	19.6	11.6	13.3	16.5	0.0
	Equal primary and English	11.8	28.3	26.4	17.3	21.3
	English w/primary to supplement	36.0	33.2	33.2	36.1	11.3
	English only	10.5	9.9	11.6	9.7	9.6
	Not taught	0.0	9.5	7.9	12.8	19.1
1	Primary only	9.4	8.3	12.1	12.2	4.0
	Primary w/English to supplement	8.1	5.6	5.6	3.9	0.0
	Equal primary and English	10.0	7.1	8.3	10.5	1.7
	English w/primary to supplement	46.8	46.5	43.4	32.7	9.9
	English only	23.9	32.2	30.4	21.5	10.3
	Not taught	1.8	0.2	0.2	19.3	74.1
2	Primary only	10.1	6.3	6.9	9.1	3.1
	Primary w/English to supplement	6.6	8.1	7.5	9.5	0.3
	Equal primary and English	7.9	1.7	1.7	2.4	0.0
	English w/primary to supplement	41.9	45.8	45.7	40.0	15.2
	English only	33.5	38.2	33.6	21.1	9.7
	Not taught	0.0	0.0	4.7	17.9	71.7
3	Primary only	4.1	9.1	4.4	3.3	0.4
	Primary w/English to supplement	11.0	9.5	13.2	18.4	0.0
	Equal primary and English	12.2	5.1	4.5	5.7	0.0
	English w/primary to supplement	27.4	27.9	30.3	24.7	5.2
	English only	45.3	48.5	46.0	30.1	24.5
	Not taught	0.0	0.0	1.7	17.8	69.9

Table 188 (continued)

Gr.	Languages	Math %	Science %	Subject Social Studies %	Ethnic Heritage %	Other %
4	Primary only	0.0	0.0	0.0	0.0	0.0
	Primary w/English to supplement	0.0	0.0	0.0	0.0	0.0
	Equal primary and English	1.8	1.8	0.0	1.8	0.0
	English w/primary to supplement	40.9	38.9	41.7	48.7	8.0
	English only	57.3	59.3	58.3	19.1	39.8
	Not taught	0.0	0.0	0.0	30.4	52.2

ERIC

Table 189

Percentage of Late-Exit Program Students
Receiving Instruction in Language(s), by Grade

Gr.	Languages	Math %	Science %	Subject Social Studies %	Ethnic Heritage %	Other %
K	Primary only	59.1	61.5	54.1	54.7	28.5
	Primary w/English to supplement	22.2	25.9	33.6	30.0	14.4
	Equal primary and English	10.2	7.8	7.5	6.3	13.2
	English w/primary to supplement	8.0	1.7	1.7	0.7	1.9
	English only	C.3	3.2	3.2	0.3	0.3
	Not taught	0.2	0.0	0.0	8.1	41.8
1	Primary only	53.7	44.9	44.9	47.2	26.1
	Primary w/English to supplement	26.0	26.9	23.6	16.5	7.6
	Equal primary and English	8.1	2.2	7.6	12.4	1.4
	English w/primary to supplement	7.8	15.5	13.3	6.9	6.7
	English only	2.6	2.1	2.1	0.7	2.5
	Not taught	1.8	8.5	8.5	16.3	55.6
2	Primary only	34.5	34.6	46.6	61.0	18.8
	Primary w/English to supplement	39.1	15.4	16.5	3.1	0.0
	Equal primary and English	14.6	16.8	25.6	14.5	0.0
	English w/primary to supplement	5.2	23.1	6.9	5.7	3.6
	English only	6.5	8.1	2.4	2.0	1.6
	Not taught	0.0	2.0	2.0	13.8	76.0
3	Primary only	6.6	14.2	24.8	29.1	3.8
	Primary w/English to supplement	21.4	18.7	11.9	8.9	10.8
	Equal primary and English	32.9	29.3	28.7	30.6	23.3
	English w/primary to supplement	29.8	23.1	24.1	15.7	11.0
	English only	6.6	13.8	8.5	3.4	13.7
	Not taught	2.8	0.9	2.1	12.4	37.5

Table 189 (continued)

		Subject				
Gr.	Languages	Math %	Science %	Social Studies %	Ethnic Heritage %	Other %
4	Primary only	1.2	5.9	8.7	16.7	16.5
	Primary w/English to supplement	31.4	32.6	22.0	28.0	0.0
	Equal primary and English	37.9	22.2	34.5	17.4	13.4
	English w/primary to supplement	15.1	19.0	21.9	15.5	2.8
	English only	13.1	17.5	9.3	6.1	9.9
	Not taught	1.4	2.8	3.7	16.2	57.4
5	Primary only	5.7	6.5	6.5	17.1	3.3
	Primary w/English to supplement	17.1	30.1	27.4	33.1	6.5
	Equal primary and English	46.9	28.2	39.6	27.4	1.3
	English w/primary to supplement	18.1	12.5	5.1	9.0	7.5
	English only	12.2	22.7	21.4	13.5	8.1
	Not taught	0.0	0.0	0.0	0.0	73.4
6	Primary only	1.4	2.8	2.8	9.6	1.4
	Primary w/English to supplement	0.0	10.3	1.4	0.0	0.0
	Equal primary and English	53.1	44.8	50.9	59.0	0.0
	English w/primary to supplement	19.9	8.4	9.8	18.3	12.7
	English only	24.2	33.7	35.1	5.0	27.6
	Not taught	1.4	0.0	0.0	8.2	58.4

Conclusions

Over five thousand student data sheets were collected from students who participated in the study. The majority of these students are of low socioeconomic status. A considerable percentage of them are considered by their teachers to be hindered in learning by lack of English proficiency, underdeveloped cognitive skills and lack of parental involvement.

Almost all students are assigned homework and most complete their homework at least half of the time.

Teachers rated their students' proficiency levels in English and Spanish. Generally, more immersion strategy students in all grades are expected to show skills in English than children in other programs and more immersion strategy students in the lower grades receive teacher ratings of "very good" on their English skills than students at lower grades in the other programs. A fair percentage of late-exit students in the higher grades also are rated as "very good". Most immersion strategy teachers do not expect their students to show proficiency in Spanish, though early-exit teachers generally expect some proficiency in Spanish, rating the majority as "good" on these skills. Late-exit teachers expect most of their students to show proficiency in Spanish, generally rating them as "fair" or "good". On the whole, students are expected to show some proficiency in computation, number concepts and verbal problem solving.

Overall, immersion strategy students receive somewhat more instruction in English language arts than do early-exit or late-exit students, but they receive almost no instruction in Spanish language arts. Across programs, teachers allow approximately equal instructional time in math, science and social studies.

While students in all three programs receive some instruction in their primary language, late-exit students receive the greatest proportions of primary language instruction, though use of English and primary language for instruction varies across programs and grades.

In sum, with a few minor exceptions, teacher reports of student behavior and performance is consistent with the program models.

VI. CONCLUSIONS

Introduction

The purpose of this chapter is to summarize the salient findings presented in earlier chapters, relate them to the main study questions, and discuss their implications for instruction and policy. Specifically, this chapter addresses the following questions: Do the three programs under study represent the three instructional models of interest, i.e., English immersion strategy, early-exit, and late-exit transitional bilingual education programs? Are the students, teachers, and administrators comparable among the three programs? What recommendations can be made regarding the learning and instruction of language minority students from these data? What do these data suggest for policy makers?

Instructional Treatment

As noted in Chapter I, if, as some language theorists posit, there are basic processes that are inherent and critical to language learning (e.g., comprehensible input, comprehensible output, and negotiation of meaning), they should be present in any successful language development program. Consequently, the major differences that would be anticipated between the three instructional models under study would be the language or languages used for instruction, how much of each language is used, and the grades in which they are used. Thus, in an immersion strategy program model, English would be used almost exclusively across all grade levels. While Spanish and English are both used in the early-exit program model, English would be the primary language of instruction, with its use rapidly increasing so that it is used almost exclusively by second grade. Students in both immersion strategy and early-exit program models would be mainstreamed at the end of the second grade. In contrast, in the late-exit program model, Spanish would be used for instruction a minimum of 40% of the time through grade six. Late-exit students who are reclassified would not be mainstreamed until grade seven. Thus the distinguishing characteristics of these instructional models are the languages used, the amount of time they are used, and the length of time students receive instruction in these languages.

Language of Instruction

Do the three study programs represent three distinct instructional models?

Yes. Data show that as grade level increases within each program — immersion strategy (grades kindergarten through four), early-exit (grades kindergarten through four), and late-exit (grades kindergarten through six) — the programs differ in the language(s) they use for instruction. Moreover, with one exception, these differences are consistent with their respective program models. English is used almost exclusively (94.3% to 98.6%) in all immersion strategy classrooms. Early-exit teachers use English approximately two-thirds of the time in kindergarten and first grade, subsequently increasing its use to approximately three-fourths in grade two, over three-fourths in grade three, and almost always in grade four. With one exception, the use of English in late-exit programs is consistent with the instructional model. Among late-exit programs, English is used sparingly in kindergarten (<10%), one-third of the time in first and second grades, about half the time in third grade, and approximately 60% in fourth grade. Contrary to its instructional model, English is used almost two-thirds of the time in grade five, and over 80% in grade six. While this pattern documents that the use of English by late-exit teachers differs markedly from that of immersion strategy and early-exit teachers, the data also suggest that late-exit classrooms in grades five and six deviate from the instructional model by exceeding the specified use of English. However, closer inspection of the data at the district level reveals that two of the three late-exit districts are using English less than 60% of the time in grade five, and English approximately two-thirds of the time in grade six. The third late-exit district uses English almost exclusively in grade five (92%) and in grade six (94%), making it more similar to an early-exit than a late-exit program.

In sum, classroom observations of teachers and students establish that the immersion strategy, early-exit, and late-exit transitional programs in this study differ with respect to the languages used for instruction. These differences are consistent with the models described in the study design for immersion strategy and early-exit programs, and with the exception of one late-exit site, they also are consistent with the late-exit program through grade five. The exception is one late-exit site that more closely resembles an early-exit instructional model than a late-exit model.

Reclassification

Do the three programs differ in the rates with which students are reclassified from LEP to FEP?

Yes. From another perspective, reclassification from limited-English-proficient to fluent-English-proficient is another indicator of program success. Contrary to expectations, by the end of four years in their respective programs (i.e., third grade), proportionately more early-exit students (72%) are reclassified than are immersion strategy students (66%). Consistent with its instructional model, while half of the LEP students in late-exit programs are reclassified by the end of third grade, almost four-fifths are reclassified by the end of sixth grade.

Mainstreaming

Do LEP students stay in each program about the same length of time?

Yes. Contrary to expectations, the amount of time language minority students remain in immersion strategy and early-exit programs is not consistent with their respective instructional models. Both immersion strategy and early-exit models call for early mainstreaming, i.e., within two to three years. This clearly does not occur. After entry into kindergarten, three-fourths of the immersion strategy and over four-fifths of the early-exit students are not mainstreamed after four years in their respective instructional programs. That is, after four years, school personnel judge that students in both programs do not have the necessary skills to be exited and work successfully in an English-only mainstream classroom. Students tend to be kept within the instructional program, even those who have been reclassified as FEP. Consistent with the instructional model, late-exit students are not mainstreamed until the seventh grade. In sum, at least for the first four years of program participation (i.e., grades kindergarten through four), students tend not to be mainstreamed in any of the three programs.

Instructional Strategies/Quality of Instruction

Differences in the three instructional models suggest that teachers in immersion strategy classrooms would have different patterns of language from those in bilingual classrooms. Immersion strategy programs, with their greater reliance on English, in theory are more likely to tailor the language to second language speakers. Bilingual programs, which in theory afford students and teachers greater opportunities for the use of L1 (Spanish in this case), should provide more complex language and instruction when L1 (i.e., student's primary language) is used for instruction than when L2 (i.e., English) is used for instruction.

Are the instructional strategies comparable among immersion strategy, early-exit, and late-exit programs?

Yes. Notwithstanding the program differences predicted and confirmed above, as anticipated in Chapter I, there are more similarities than differences in instructional strategies among the three programs. Instructional strategies are defined by teacher language behavior. Teachers are positive and supportive of their students through the feedback they provide to them. The most common types of statements made by teachers are the same for all three programs and grade levels: explanation, question, command and feedback. Additionally, teachers tend to make the same types of statements in both languages. That is, teachers tend not to use English and Spanish differently. Across all programs, teacher questions typically are low-level requests for simple information recall. Classroom activities tend to be teacher-directed. Nonetheless, instructional strategies are positive and supportive of students in each program.

Do classrooms in the three instructional programs provide an ideal language learning environment?

No. Without exception, across grade levels within and between the three instructional programs, students are limited in their opportunities to produce language and in their opportunities to produce more complex language. This is evident from observations of teachers doing most of the talking in classrooms, making about twice as many utterances as do students. Students produce language only when they are working directly with a teacher, and then only in response to teacher initiations. There are two major concerns from these findings. The first concern is that in over half of the interactions that teachers have with students, students do not produce any language (i.e., non-verbal responses such as listening, gesturing, etc.). Second, when students do respond, typically they are providing simple information recall. Rather than being provided with the opportunity to generate original statements, students are asked to provide simple discrete close-ended or patterned responses. Not only does this pattern of teacher/student interaction limit a student's opportunity to create and manipulate language freely, but it also limits the student's ability to engage in more complex learning (i.e., higher order thinking skills).

In sum, teachers in all three programs do not teach language or higher order cognitive skills effectively. Teachers in all three programs offer a passive language learning environment, limiting student opportunities to produce language and to develop more complex language and conceptual skills.

Do teachers speak differently to limited-English-proficient (LEP) students than to fluent-English-proficient (FEP) and/or English-only (EO) students?

No. Teachers across programs and grades tend to say the same things to LEP, FEP, and EO students when they are separated by these language classifications. However, teachers speak quite differently to students when student groups are comprised of LEPs, FEPs, and/or EOs than when groups are separated by language classification levels. Teachers tend to explain, model and monitor more often when speaking to mixed groups of students (thereby ensuring LEP students have enough information and direction to participate in the lesson) than when the students are separated by language classification levels. Teachers also ask fewer questions, give fewer commands, and provide less feedback when students are mixed by language status than when they are separated. Thus, teachers in all three programs exhibit some sensitivity to the special needs of LEP students when they are mixed with FEP and/or EO students.

Are students engaged in their assigned tasks?

Yes. Consistently across programs, grades, and language classification levels, students exhibit a high level of task engagement ($\geq$86.5%). This suggests that teachers across all programs are equally effective in providing lessons that meet the needs of their students.

Do teachers shelter their use of English through realia?

Yes, albeit limited. There is some evidence of teachers' attempts to adjust their language behavior through realia. However, the use of realia appears to be limited to kindergarten classrooms by immersion strategy and early-exit kindergarten teachers. Late-exit teachers tend not to use realia at any grade level.

Do the three programs differ in the complexity of language, content, or context of utterances?

Overall, no. Complexity, as defined as students responding with free responses, use of less repetition and drill, and greater use of opportunities for discussion, differs little or not at all between programs. Program differences in the complexity of language are noted in one area, with proportionately more late-exit students asking questions than immersion strategy students.

Are the three programs comparable in the amount of English language arts instruction provided?

No. Immersion strategy teachers tend to devote more time to English language arts than early-exit or late-exit teachers. However, early-exit and late-exit teachers increase the proportion of time they devote to English language arts as grade level increases. Moreover, when one considers the Spanish language arts instruction provided by early-exit and late-exit teachers, the total amount of instructional time spent on language arts, English and/or Spanish, is comparable across the three programs.

Are the three programs comparable in the total amount of instruction provided per day by content area?

Yes. Each program allocates approximately the same amount of instructional time in reading, language arts (collapsing instruction in English and Spanish), math, and social studies.

Are the three programs comparable in the assignment and correction of homework?

No. Consistently, late-exit teachers assign and grade homework with greater frequency than do immersion strategy or early-exit teachers.

ERIC

Do classroom instructional activities vary by program and grade?

Yes, by grade, but only slightly. Typically, students in all programs are provided with seatwork and discussion activities, with low to moderate involvement in drillwork, listening, interim or other activities. The frequency of these passive learning activities and discussion activities wherein students do not produce much language (and that which is produced is typically information recall) provides students with a less than ideal language learning environment.

Did observations of teacher/student interactions take place during comparable settings across the three programs?

Overall, yes. Observations were made in each program during each content area (i.e., reading, language arts, math, social studies, non-academic such as art or music, other academic such as science or computer lab, or procedures such as making a transition from one subject area to another). Only minor grade level differences were noted between programs.

Is the involvement of the parents of target students in the schooling of their children comparable across the three programs?

No. Late-exit parents are more aware that their children have homework and ensure that it is completed than either immersion strategy or early-exit parents. Except for kindergarten, teachers across grades and programs always assign homework. Yet only three-fourths (75%) of early-exit and four-fifths (83%) of immersion strategy parents report that their children have homework. In contrast, almost all (98%) late-exit parents are aware that their children have homework. Finally, proportionately more late-exit parents monitor the completion or help their children complete their homework (93.3%) than either immersion strategy or early-exit parents (68.2% and 70.6%, respectively).

Do the instructional programs reflect the educational goals of the target students' parents?

No. The majority of parents across programs want their children to know Spanish and English equally well. However, there was substantial variability among parents across and within programs as to how these goals might be realized. These differences reflect the different approaches used by each of the three instructional programs as communicated to parents as being in the best interests of their children.

Conclusion

In conclusion, the three programs in this study represent three distinct instructional models, differing primarily in the amount of English and Spanish used for instruction. In all, the language use patterns are consistent with those of their respective instructional models. However, immersion strategy and early-exit programs tend not to reclassify or mainstream their students early, but keep them in their respective programs for at least five years (i.e., through grade four). Consistent with the instructional model, late-exit students are reclassified (but not mainstreamed) more slowly than either immersion strategy or early-exit students.

The data show that the programs in this study are more similar than different in the instructional strategies used. Regardless of the language used or the language classification of students, the basic instructional paradigm is explanation, question, command and feedback. Of concern is that the instructional strategies used by teachers in all three programs make for a passive language learning environment, limiting students' opportunities to develop more complex language and conceptual skills.

The data suggest that the three programs are comparable, with two exceptions, with respect to the quality of instruction provided. Various indicators of instructional quality (i.e., engaged academic time, use of realia, complexity of language, content, or context of utterances) also reveal little or no program differences. However, two important indicators of instructional quality suggest an advantage of late-exit programs. First, late-exit teachers assign and correct homework more often than either immersion strategy or early-exit teachers. Second, research in the role of parents in the education of their children, especially for language minority parents, documents the importance of parental involvement in their children's learning (Ramirez, Douglas, and Vargas, 1989). Thus, parent involvement can be used as an indicator of program quality. Language minority parents whose children are in early-exit or late-exit programs are involved more actively in the schooling of their children than parents of children in immersion strategy programs. Furthermore, across programs, the majority of language minority parents want their children to develop equally their primary and English language skills. From another perspective, the late-exit program again better represents the educational preferences of language minority parents and is reflected in their greater satisfaction with the schooling of their children than either immersion strategy or early-exit parents.

Student Characteristics

Are target students comparable between programs with respect to their home backgrounds?

Overall, yes. Target students come from families with two adults and three children. While most target students were born in the United States, over three-fourths of their parents were not. Almost three-fourths of the families are of low income (<$15,000/year). However, proportionately more late-exit families are of the lowest income levels than are early-exit or immersion strategy families. While roughly one-third of the target students attended preschool, proportionately more immersion strategy students than either early-exit or late-exit students attended preschools where only English was spoken.

Is the home environment support for English and Spanish language use among target students comparable between programs?

Yes. While most communication with parents and children in the home is in Spanish, children use more English with siblings. Approximately half of the households subscribe to English language newspapers and half to Spanish language newspapers. Almost all families have a television set, and children watch mostly English programming. Approximately two-thirds of other media sources (i.e., radio, tape, and record player) provide language in English and one-third in Spanish. Most target students play with friends who use both English and Spanish, and they live in neighborhoods where Spanish is spoken at least half of the time. These characteristics are consistent across all programs.

In sum, with one exception, students across programs have comparable backgrounds. The exception is that proportionately more late-exit students are from the lowest income levels than are immersion strategy or early-exit students.

Teacher/Classroom Characteristics

Teacher Characteristics

Are teachers comparable between programs with respect to their language proficiency, ethnicity, training, and attitudes towards the education of LEP students?

No. Late-exit teachers tend to have backgrounds similar to their students, are sufficiently fluent in Spanish to teach in it, and have advanced training in meeting the needs of language minority students. In contrast, immersion strategy and early-exit teachers generally are not Hispanic, are not sufficiently proficient in Spanish to teach in it, and do not have as much advanced training as do late-exit teachers. Moreover, teachers in each program differ markedly in their attitudes on how limited-English-proficient students should be taught, essenti-

ERIC

ally concurring with the underlying rationale for their respective instructional models.

Classroom Characteristics

Are classrooms comparable between programs with respect to class size?

Yes. Classes across programs range in size from a mean of 21.8 to 27.1 students per class.

Are classrooms comparable between programs with respect to classroom composition (i.e., mix of LEP, FEP, and EO)?

No. Late-exit classrooms have noticeably fewer EO students than immersion strategy or early-exit classrooms. In each program, the proportion of FEP students increases and the proportion of LEP students decreases as grade level increases (suggesting a modicum of success among each program).

Are student grouping patterns comparable between programs?

No. Immersion strategy and early-exit teachers tend to use large-sized (>10 students) and small-sized (2-5 students) groups. In contrast, late-exit teachers more characteristically use large- and medium-sized (6-10 students) groups.

School/District Characteristics

The success of any instructional program in large measure also reflects the administrative context in which it is implemented. Thus the knowledge and support of site and district level administrators as well as the available resources of the school, district, and community may act to support or weaken the implementation of an instructional program.

Project Administrators

Are project administrators comparable between programs regarding their background and training?

No. All project administrators of late-exit programs are Hispanic, have Spanish as their first language, and have a master's degree or higher; most are certified bilingual or ESL teachers and have direct bilingual/ESL teaching experience. In contrast, immersion strategy administrators tend to be white, speak Spanish as their second language, and have approximately half of the bilingual/ESL training and experience that late-exit project administrators have. Early-exit project administrators fall in between late-exit and immersion strategy project administrators on these characteristics.

Site Administrators (Principals)

Are site administrators comparable between programs regarding their personal and professional background?

No. Proportionately more early-exit and late-exit principals have backgrounds similar to that of their students, have training in working with language minority children, and have practical experience in teaching these children than do principals in immersion strategy programs.

Are site administrators comparable between programs in their understanding of their respective instructional programs?

Yes. Across programs, site administrators appear to understand the theoretical and methodological concepts underlying their respective programs. However, principals are not as well prepared in terms of training and experience as are project directors. More importantly, principals' attitudes and opinions regarding the education of language minority children do not always agree with those of their respective project administrators. This difference may serve to constrain full implementation of their respective programs. This finding is consistent across all programs.

School Characteristics

Is the student body in each of the schools comparable between programs?

No. While all school sites report substantial numbers of their student body as language minority (one-third to three-fourths) and low income, program differences are evident. The availability of EO students as role models is important for development of English language skills. Proportionately more of the students at early-exit only and immersion strategy/early-exit school sites than in late-exit or immersion strategy only school sites are language minority students. Immersion strategy principals report that proportionately more of their student body comes from homes with the lowest income levels as compared to other school sites.

Do schools between programs have comparable procedures for assessing the language proficiency skills among language minority students?

Yes. Most schools use more than one procedure for assessing language proficiency for initial classification and for reclassification. In most cases, students classified as LEP automatically receive language services unless their parents request otherwise. In most instances, when immersion strategy and early-exit LEP students are reclassified as FEP, they are expected to be transferred to mainstream, English-only classrooms.

Are schools comparable between programs in the extent to which they mix their LEP, FEP, and EO students?

No. Schools having both an immersion strategy and an early-exit program provide the most heterogeneous mix of students. Less than one-half of immersion strategy only and two-fifths of late-exit school sites do so. However, in all programs, mixing of students typically

ERIC

tends to occur in grades four through six rather than in kindergarten through grade three. Waiting to mix LEP students with EO students is contrary to second language acquisition research documenting how interaction with native speakers greatly facilitates second language acquisition. In sum, primary school teachers are limiting the language learning opportunities for LEP students in the early primary grades by separating them from EO students (assuming the availability of EO students) until the fourth grade.

<u>Are schools comparable between programs as to the extent that their instructional programs are articulated and coordinated?</u>

Yes. Almost all schools report having written curricula for all content areas, and more than half of the early-exit and late-exit school sites have such curricula for Spanish language arts and math. Student progress is monitored by principals and teachers, and relayed to staff through meetings and to parents through conferences and bulletins. Parent involvement tends to be limited to PTAs and parent advisory committees.

<u>District Context</u>

<u>Are school districts comparable between programs in the proportion of their student body identified as LEP?</u>

No. Collapsing across grades, early-exit/immersion strategy districts tend to have more of their student body identified as LEP than do late-exit districts (27.5% to 38.3%, and 14.3% to 21.7%, respectively). However, with one exception, all districts in the three programs note a consistent decline in the proportion of LEP students as grade level increases from kindergarten to sixth grade, suggesting a modicum of success in each program.

Are school districts comparable between programs in the proportion of students mainstreamed?

Yes. Contrary to their instructional models and district policies, immersion strategy and early-exit school districts do not mainstream their students at a greater rate than do late-exit districts. That is, immersion strategy and early-exit districts tend to keep their students in their respective programs for at least five years (i.e., through grade four, reflecting the limits of the study rather than of the programs).

Implications

- Contrary to the objectives of immersion strategy and early-exit programs, most students remain in these programs much longer than expected. It is clear that immersion strategy and early-exit teachers believe that the majority of limited-English-proficient students would be better off if they remain in the programs for more than three years. The limited evidence from this study suggests that limited-English-proficient students may need prolonged assistance if they are to succeed in English-only mainstream classrooms.

In spite of the strong support for early mainstreaming among immersion strategy and early-exit staff (instructional and administrative), while two-thirds to three-fourths of language minority students are reclassified by their teachers as FEP by the end of third grade, fewer than one-fourth are mainstreamed by grade four. This suggests that, at a practical level, immersion strategy and early-exit teachers feel that most LEP students need five or more years of special instructional support prior to mainstreaming. Neither program is able to provide a "quick fix." Federal guidelines of services to language minority students should be reviewed to reflect the professional judgments made by those working directly with these students.

o **There is a need to improve the quality of training programs for teachers serving language minority students both at the university and school district levels, so that they can provide a more active learning environment for language and cognitive skill development. Effective training models do exist which can help teachers provide a more active learning environment for language and cognitive skill development. Efforts should be made to disseminate this information and support implementation of these models.**

Across districts, states, and programs, while teachers who have been specially trained to meet the instructional needs of language minority students (i.e., holding a bilingual or ESL teaching credential) know what they are supposed to do, they do not do so. That is, they provide a passive language learning environment. Contrary to the training they have received on current theory and research on language development and second language acquisition, these specially trained teachers limit students' opportunities to produce language and to develop more complex language and conceptual skills, a finding that is consistent with studies of mainstream English-only teachers. Why? There are two possible explanations, both highlighting basic problems with current teacher training efforts.

First, teachers may be simply reflecting the underlying assumptions of the training that they received in teacher preparation. The data are consistent with the assertion by Cummins (1986) that teacher training in the United States is basically "transmission oriented" rather than "reciprocal/interaction oriented." That is, teachers may be "lectured" during training on the importance of active learning strategies rather than experiencing a modeling of alternative active learning strategies.

Secondly, teacher training efforts may not incorporate all of the essential components of an effective training program: theory, modeling, practice, and on-going coaching (Calderon, 1982; Joyce & Showers, 1983). Each component is important and should not be overlooked. An understanding of theory provides the teacher with a conceptual framework of the rationale for a given instructional approach. Modeling by the trainer allows the teachers in training to observe how the theory is operationalized. Practice allows the teachers in training to demonstrate their understanding of what they have learned. Our data suggest that these three components are minimally present in teacher training efforts in so far as teachers are able to describe an ideal language learning environment. However, it is on-going coaching (observing others and being observed) in a classroom setting that crystallizes this learning and makes it possible for the teachers in training to incorporate the instructional approach into their teacher repertoire (Joyce & Showers, 1982; Berliner, 1982; Wood, McQuarrie, and Thompson, 1982). As the average teacher in this study failed to provide an optimal language learning environment, one might hypothesize that this reflects a lack of sustained teacher training support (i.e., on-going coaching).

It is recommended that federal efforts encourage and examine teacher training efforts to include all the requisite components of a comprehensive training model.

o Parental involvement appears to be greatest in the late-exit program. This suggests that schools serving limited-English-proficient students should explore how they might use the home language of their students to engage parents in the schooling of their children.

By virtue of their children receiving some of their work in their primary language, parents are able to assist their children in their work more directly and more frequently. Using the home language for instruction appears to facilitate parent participation. As the home language is used more for a longer period of time in the late-exit program than in either the immersion strategy or early-exit programs, parents are provided with more opportunities to support children's learning over a longer grade span. This greater involvement may also partly explain the greater satisfaction of their children's education by late-exit parents than reported by immersion strategy or early-exit parents.

REFERENCES

Baker, K., & de Kanter, A. (1981). Effectiveness of bilingual education: A review of the literature. Washington, D.C.: U.S. Department of Education, Office of Planning, Budget, and Evaluation.

Baker, K., & de Kanter, A. (1983). Bilingual education: A reappraisal of federal policy. Lexington, MA: Lexington Books.

Barik, H., & Swain, M. (1976). A Canadian experiment in bilingual education: The Peel Study. Foreign Language Annals, 9, 465-479.

Barik, H., & Swain, M. (1977). Report to the Elgin County Board of Education Re: Evaluation of the 1976-77 partial French immersion program in grades 5-7. Toronto: Ontario Institute for Studies in Education. Mimeo.

Barik, H., & Swain, M. (1978). Evaluation of a bilingual education program in Canada: The Elgin Study through grade six. Bulletin CILA, 27, 31-58.

Barik, H., Swain, M., & Gaudino, V. (1976). A Canadian experiment in bilingual education in the senior grades: The Peel Study through grade 10. International Review of Applied Psychology, 25, 99-113.

Barik, H., Swain, M., & Nwanunobi, E. (1977). English-French bilingual education: The Elgin Study through grade five. Canadian Modern Language Review, 33, 459-475.

Berliner, D. (1982). On improving teacher effectiveness: A conversation with David Berliner. Educational Leadership, 12-16.

Bruck, M. (1978). The suitability of early French immersion programs for the language disabled child. Canadian Modern Language Review, 34, 884-887.

Bruck, M. (1979). Problems in early French immersion programs. In B. Mlacak and E. Isabelle (Eds.), So you want your child to learn French. Ottawa: Canadian Parents for French, 42-47.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. Psychological Bulletin, 101, 147-158.

Bryk, A. S., Raudenbush, S. W., Seltzer, M., & Congdon, R. J. (1986). An introduction to HLM: Computer program and user's guide. Chicago: University of Chicago.

ERIC

Calderon, M. (1982). Second Language Acquisition And Organizational Change Through An Effective Training System. San Diego, CA: Bilingual Education Service Center.

Campbell, R. (1984). The immersion education approach to foreign language teaching. In Studies on immersion education: A collection for United States Educators. Sacramento: California State Department of Education, 114-143.

Cazden, C. (1984). Effective instructional practices in bilingual education. Unpublished manuscript, Harvard University.

Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learner's errors. Language Learning, 27(1), 29-46.

Chaudron, C. (1985a). Transcription directions for SRA project. Mountain View, CA: SRA Technologies/Second Language Research.

Chaudron, C. (1985b). Instructions for verifying transcripts. Mountain View, CA: SRA Technologies/Second Language Research.

Chaudron, C. (1985c). Coding procedures and guidelines for SRA project. Mountain View, CA: SRA Technologies/Second Language Research.

Chaudron, C. (1985d). Directions for tallying transcripts. Mountain View, CA: SRA Technologies/Second Language Research.

Cohen, A. (1975). A sociolinguistic approach to bilingual education. Rowley, MA: Newbury House.

Collier, V. P. (1989). How long? A synthesis of research on academic achievement in a second language. TESOL Quarterly, 23(3), 509-531.

Crespo, O. I. (1986). Longitudinal Study of Immersion Programs for Language-Minority Children. Washington, D.C.: Pelavin Associates.

Cummins, J. (1975). Cognitive factors associated with intermediate levels of bilingual skills. Dublin: St. Patrick's College, Educational Research Centre.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. Working Papers on Bilingualism, 9, 1-43.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In Schooling and language minority students: A theoretical framework. Los Angeles: California State University; Evaluation, Dissemination and Assessment Center, pp. 3-49.

ERIC

Cummins, J. (1983). Functional language proficiency in context: Classroom participation as an interactive process. In W. Tikunoff (Ed.), Compatibility of the SBIF features with other research: An instruction for LEP students. San Francisco: Far West Laboratory for Educational Research and Development.

Cummins, J. (1984). Bilingualism and special education: Issues in assessment and pedagogy. Clevedon, England: Multilingual Matters. (Co-published in the United States by College-Hill Press, San Diego)

Cummins, J. (1986). Empowering minority students: A framework for intervention. Harvard Educational Review, 56(1), 18-36.

Cummins, J., & Swain, M. (1986). Bilingualism in education: Aspects of theory, research and practice. London: Longman.

Cziko, G., Lambert, W., & Gutter, R. (1979). French immersion programs and students' social attitudes: A multidimensional investigation. Working Papers on Bilingualism, 19, 13-28.

Dulay, H., & Burt, M. (1978). Why bilingual education? A summary of research findings (2nd ed.). San Francisco: Bloomsbury West.

Dulay, H., & Burt, M. (1979, September). Bilingual education: A close look at its effects. NABE News, 3(1).

Dunkin, M. J., & Biddle, B. J. (1974). The study of teaching. New York: Holt, Rinehart & Wilson.

Epstein, N. (1977). Language, ethnicity, and the schools: Policy alternatives for bilingual-bicultural education. Washington, D.C.: George Washington University, Institute for Educational Leadership.

Fisher, C., & Guthrie, L. (1983). Executive summary: The significant bilingual instructional features study, Part II. San Francisco: Far West Laboratory for Educational Research and Development.

Fishman, J., & Lovas, J. (1970). Bilingual education in sociolinguistic perspective. TESOL Quarterly, 4, 215-222.

Gaies, S. J. (1977). The nature of linguistic input in formal second language learning: Linguistic and communicative strategies in ESL teachers' classroom language. In H. D. Brown, C. A. Yorio, & R. H. Crymes, (Eds.), On TESOL, '77. Teaching and learning English as a second language: Trends in research and practice. Washington, D.C.: TESOL.

Genesee, F. (1974). An evaluation of the English writing skills of students in French immersion programs. Montreal: Protestant School Board of Greater Montreal. Mimeo.

Genesee, F. (1976). The role of intelligence in second language learning. Language Learning, 26, 267-280.

Genesee, F. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. Language Learning, 27, 29-46.

Genesee, F. (1978). A longitudinal evaluation of an early immersion school program. Canadian Journal of Education, 3(4), 31-50.

Genesee, F. (1983). The bilingual education of majority language children: The immersion experiments in review. Applied Psycholinguistics, 4, 1-46.

Genesee, F. (1984). Historical and theoretical foundations of immersion education. In Studies on immersion education: A collection for United States educators. Sacramento: California State Department of Education, 32-57.

Genesee, F. (1985). Second language learning through immersion: A review of U.S. programs. Review of Educational Research, 55(4), 541-561.

Genesee, F. (1987). Learning through two languages: Studies of immersion and bilingual education. Cambridge, MA: Newbury House.

Genesee, F., & Lambert, W. (1983). Trilingual education for majority language children. Child Development, 54, 105-114.

Genesee, F., Polich, E., & Stanley, M. (1977). An experimental French immersion program at the secondary school level: 1969-1974. Canadian Modern Language Review, 33, 318-332.

Genesee, F., Tucker, G. R., & Lambert, W. (1975). Communication skills of bilingual children. Child Development, 46, 1010-1014.

Goldstein, H. (1986). Efficient statistical modeling of longitudinal data. Annals of Human Biology, 13(2), 129-141.

Hamayan, E., & Tucker, G. R. (1980). Language input in the bilingual classroom and its relationship to second language achievement. TESOL Quarterly, 14, 453-468.

Harley, B. (1979). French gender rules in the speech of English-dominant, French-dominant and monolingual French-speaking children. Working Papers on Bilingualism, 19, 129-156.

Harley, B. (1982). Age-related differences in the acquisition of the French verb system by Anglophone students in French immersion programs. Unpublished doctoral dissertation, University of Toronto.

ERIC

Hernandez-Chavez, E. (1984). The inadequacy of English immersion education as an educational approach for language minority students in the United States. In Studies on immersion education: A collection for United States education. Sacramento: California State Department of Education, 144-183.

Joyce, B., & Showers, B. (1982). The coaching of teaching. Educational Leadership, 4-10.

Joyce, B., & Showers, B. (1983). Power in Staff Development through Research in Training. Oregon: Association for Supervision and Curriculum Development.

Judd, C., & Kenny, D. (1981). Estimating the effects of socio interventions. Cambridge: Cambridge University Press, 130.

Kjolseth, R. (1972). Bilingual education program in the United States: For assimilation of pluralism? In B. Spolsky (Ed.), The language education of minority children. Rowley, MA: Newbury House, 94-121.

Krashen, S. (1981). Bilingual education and second language acquisition theory. In Schooling and language minority students: A theoretical framework. Los Angeles: Evaluation, Dissemination and Assessment Center, California State University.

Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. Biometriks, 38, 963-974.

Lambert, W. (1984). An overview of issues in immersion education. In Studies on immersion education: A collection for United States educators. Sacramento: California State Department of Education, 8-30.

Lambert, W., & Tucker, G. R. (1972). Bilingual education of children: The St. Lambert experiment. Rowley, MA: Newbury House.

Lapkin, S. (1982). The English writing skills of French immersion pupils at grade five. Canadian Modern Language Review, 39, 24-33.

Lapkin, S., & Cummins, J. (1984). Canadian French immersion education: Current administrative and instructional practices. In Studies on immersion education: A collection for United States educators. Sacramento: California State Department of Education, 58-86.

Legarreta, D. (1977). Language use in bilingual classroom. TESOL Quarterly, 11, 9-16.

Long, M. (1983). Inside the "Black Box": Methodological issues in classroom research on language learning. In H. Seliger & M. Long (Eds.), Classroom oriented research in second language acquisition. Rowley, MA: Newbury House.

Mackey, W. (1972). Bilingual education in a binational school. Rowley, MA: Newbury House.

McEachern, W. (1980). Parental decision for French immersion: A look at some influencing factors. Canadian Modern Language Review, 38, 238-246.

McGraw-Hill, Inc. (1974). Comprehensive Tests of Basic Skills, Form S, Technical Bulletin No. 1. Monterey, CA: CTB/McGraw-Hill.

McLaughlin, B. (1982). Language learning in bilingual instruction: Literature review. Santa Cruz: University of California.

Milk, R. (1980). Variation in language use patterns across different group settings in two bilingual second grade classrooms. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Morrison, F. (1981). Evaluation of the second language learning (French) programs in schools of the Ottawa and Carleton boards of education. Eighth annual report. Ottawa: Ottawa Board of Education.

Nerenz, A., & Knop, C. (1983). Allocated time, curricular content, and student engagement outcomes in the second language classroom. Canadian Modern Language Review, 39, 222-232.

Office of Bilingual Bicultural Education, California State Department of Education (1981). Annual report. Sacramento: Author.

Office of Civil Rights. (1975). Task force findings specifying remedies available for eliminating past educational practices ruled unlawful under Lau vs. Nichols. Washington, D.C.: Author.

Omaggio, A. (1982). The relationship between personalized classroom talk and teacher effectiveness ratings: Some research results. Foreign Language Annals, 15, 255-269.

Paulston, C. (1980). Bilingual education theories and issues. Rowley, MA: Newbury House.

Paulston, C. (1983). Second language acquisition in school settings. In W. Tikunoff (Ed.), Compatibility of the SBIF features with other research and instruction for LEP students. San Francisco: Far West Laboratory for Educational Research and Development.

Penfield, W., & Roberts, L. (1958). Speech and brain mechanisms. New York: Atheneum.

Plann, S. (1976). The Spanish immersion program: Towards native-live proficiency of classroom dialect?. Unpublished master's thesis, University of California at Los Angeles.

Politzer, R. (1977). Foreign language teaching and bilingual education: Implications of some recent research findings. Paper presented to the ACTFL Annual Conference, San Francisco, CA.

Ramirez, A., & Stromquist, N. (1979). ESL methodology and student language learning in bilingual elementary schools. TESOL Quarterly, 13, 145-158.

Ramirez, J. D. (1981). The relationship of adult/child ratio and group size to caregiver behavior and child activities. Unpublished doctoral dissertation, Stanford University, Stanford, CA

Ramirez, J. D., Douglas, D., & Vargas, R. (1989). Language minority parents and the school: Home-School partnerships, bridging the gap. Office of Bilingual Education, California State Department of Education.

Ramirez, J. D., & Merino, B. (1989). Classroom talk in English immersion, early-exit and late-exit transitional bilingual education programs. In Jacobson, R., & Faltis, C. (Eds.), Language Distribution in Bilingual Schooling. Clevendon, England: Multilingual Matters, Ltd.

Ramirez, J. D., Schinke-Llano, L., & Bloom, G. (1984). The characteristics and effectiveness of transitional bilingual education programs. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J. D., Wolfson, R., Tallmadge, G. K., & Merino, B. (1984). Study design of the longitudinal study of immersion programs for language-minority children. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J. D., Wolfson, R., Morales, F. (1985). Description of immersion strategy programs in the United States. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J. D., Yuen, S. D., & Ramey, D. R. (1987). Second year report: Longitudinal study of immersion, early-exit and late-exit transitional bilingual education programs for language minority children. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J. D., Yuen, S. D., & Ramey, D. R. (1988). Third year report: Longitudinal study of immersion, early-exit and late-exit transitional bilingual education programs for language-minority children. San Mateo, CA: Aguirre International.

ERIC

Raudenbush, S. W. (1988). Educational applications of hierarchical models: a review. Journal of Educational Statistics, 13, 85-116.

Raudenbush, S. W. (1988). The hierarchical linear model as a generalized multivariate linear model. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Redfield, D., & Rousseau, E. (1981). A meta-analysis of experiemental research on teacher questioning behavior. Review of Educational Research, 51(2), 237-245.

Reiss, A. J. (1961). Occupations and Social Status. New York: The Free Press of Glencoe.

Schinke-Llano, L. (1983). Foreigner talk in content classrooms. In H. Seliger & M. Long (Eds.), Classroom oriented research in second language acquisition. Rowley, MA: Newbury House.

Schinke-Llano, L., & Ramirez, J. D. (1984). Programmatic and Instructional Aspects of Language Immersion Programs. Mountain View, CA: SRA Technologies, Inc.

Schultz, J. (1975). Language use in bilingual classrooms. Paper presented at the Annual Convention of Teachers of English to Speakers of Other Languages (TESOL), Los Angeles, CA.

Spilka, I. (1976). Assessment of second language performance in immersion programs. Canadian Modern Language Review, 32, 543-561.

Stanford University. (1976). Program on Teaching Effectiveness. Stanford, CA: Stanford University.

Strenio, J. L. F., Weisberg, H. I., & Bryk, A. S. (1983). Empirical Bayes estimation of individual growth curves and their relationships to covariates. Biometrics, 39, 71-86.

Swain, M. (1975). Writing skills of grade three French immersion pupils. Working Papers on Bilingualism, 7, 1-38.

Swain, M. (1980). French immersion programs in Canada. Multiculturalism, 4(2), 3-6.

Swain, M. (1983). Bilingualism without tears. In M. Clarke and J. Handscombe (Eds.), On TESOL '82. Washington, D.C.: TESOL, 35-46.

Swain, M. (1984). A review of immersion education in Canada: Research and evaluation studies. In Studies on immersion education: A collection for United States educators. Sacramento: California State Department of Education, 87-112.

ERIC

Swain, M., & Lapkin, S. (1982). Evaluating bilingual education: A Canadian case study. Clevedon, England: Multilingual Matters.

Tallmadge, G. K. (1978). Achievement testing of disadvantaged and minority students for educational program evaluation. Monterey, CA: CTB/McGraw-Hill.

Terrell, T. (1981). The natural approach in bilingual education. In schooling and language minority students: A theoretical framework. Los Angeles: California State University, National Evaluation, Dissemination, and Assessment Center, 117-146.

Tikunoff, W. (1983a). An emerging description of successful bilingual instruction: Executive summary of Part I of the SBIF study. San Francisco: Far West Laboratory for Educational Research and Development.

Tikunoff, W. (Ed.) (1983b). Compatibility of the SBIF features with other research on instruction for LEP students. San Francisco: Far West Laboratory for Educational Research and Development.

Tikunoff, W., & Vazquez-Faria, J. (1982). Peabody Journal of Education, 59, 234-271.

Trueba, H., & Barnett-Mizrahi, C. (Eds.) (1979). Bilingual multicultural education and the professional: From theory to practice. Rowley, MA: Newbury House.

Trueba, H., & Wright, P. (1981). Journal of Multilingual and Multicultural Development, 2, 243-257.

Tucker, G. R. (1980). Implications for U.S. bilingual education: Evidence from Canadian research. Focus, 2, National Clearinghouse for Bilingual Education.

Willett, J. (1988). Questions and answers in the measurement of change. Review of Research in Education, 15, 433-516.

Willig, A. (1981, Winter/Spring). The effectiveness of bilingual education: Review of a report. NABE Journal, 6(2/3).

Willig, A. (1985). Review of Educational Research.

Wilson, B. L., Rosenholtz, J., & Rosenholtz, H. (1983). Effect of task and authority structures on student task engagement. Paper presented at annual meeting of AERA, Montreal.

Wong-Fillmore, L. (1983). Effective language use in bilingual classes. In W. Tikunoff (Ed.), Compatibility of the SBIF features with research on instruction for LEP students. San Francisco: Far West Laboratory for Educational Research and Development.

Wong-Fillmore, L., & Valadez, C. (1986). Teaching bilingual learners. In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed.). New York: Macmillan.

Wong-Fillmore, L., Ammon, P., McLaughlin, B., & Ammon, M. (1985). Learning English through bilingual instruction. Final report submitted to the National Institute of Education, Berkeley: University of California.

Wong-Fillmore, L., Ammon, P., Ammon, M., DeLucchi, K., Jensen, J., McLaughlin, B., & Strong, M. (1983). Learning language through bilingual instruction: Second year report. Submitted to the National Institute of Education, Berkeley: University of California.

Wood, F., McQuarrie, F., & Thompson, S. (1982). Practitioners and professors agree on effective staff development practices. Educational Leadership, 28-31.

Zappert, L., & Cruz, B. R. (1977). Bilingual education: An appraisal of empirical research. Berkeley, CA: Bahia, Inc.

APPENDIX A

Weighting Algorithm For Descriptive Analyses

ERIC

DIRECTORY OF TABLES

APPENDIX A

WEIGHTING ALGORITHM FOR DESCRIPTIVE ANALYSES

The number of students in each district/program combination was defined to be the number of students in that district/program who participated in the study for portions of at least two years. This group represents the students who contributed the most data.

Table A1 displays the overall proportion of students in the student data base for each district/program combination, the proportion of students in the student data base for each district/program combination who participated for portions of at least two years, and the proportions actually used to determine the weights for each instrument (the proportions were adjusted slightly to give each program equal weight). Tables A2 to A11 provide the weights for each instrument.

ERIC

Table A1

Proportion of Students in Student Data Base for each District/Program Combination

	Program								
	Immersion			Early-Exit			Late-Exit		
District	Overall	2+ yrs.	Adjusted	Overall	2+ yrs.	Adjusted	Overall	2+ yrs.	Adjusted
0:A	6.5	6.6	6.4	11.0	7.8	7.8	*	*	*
1:B	8.3	9.7	9.4	9.5	10.6	10.7	*	*	*
2:C	4.7	4.6	4.4	5.3	4.9	4.9	*	*	*
3:D	*	*	*	*	*	*	3.8	4.4	4.5
4:E	*	*	*	*	*	*	16.5	19.5	20.0
5:F	3.2	2.7	2.7	7.2	2.6	2.6	*	*	*
6:G	*	*	*	*	*	*	7.9	8.5	8.8
7:H	9.1	10.9	10.5	*	*	*	*	*	*
8:I	*	*	*	6.9	7.4	7.4	*	*	*
			33.4			33.4			33.3

Table A2

LOM/CEATM
Weights assigned to each District/Program Combination

District	Program		
	Immersion	Early-Exit	Late-Exit
A	1.074	0.671	*
B	1.197	1.243	*
C	1.026	1.138	*
D	*	*	0.608
E	*	*	1.107
F	2.227	0.899	*
G	*	*	0.902
H	0.979	*	*
I	*	1.003	*

ERIC Full Text Provided by ERIC

Table A3

Classroom Data Base
Weights assigned to each District/Program Combination

District	Program		
	Immersion	Early-Exit	Late-Exit
A	0.962	0.641	*
B	1.219	1.235	*
C	1.021	1.235	*
D	*	*	0.625
E	*	*	1.162
F	1.345	0.572	*
G	*	*	0.860
H	1.145	*	*
I	*	1.178	*

Table A4

Teacher Interview
Weights assigned to each District/Program Combination

District	Program: Immersion	Early-Exit	Late-Exit
A	0.938	0.625	*
B	1.188	1.205	*
C	0.995	1.261	*
D	*	*	0.625
E	*	*	1.209
F	1.440	0.559	*
G	*	*	0.838
H	1.187	*	*
I	*	1.217	*

ERIC

Table A5

Student Data Base
Weights assigned to each District/Program Combination

District	Program		
	Immersion	Early-Exit	Late-Exit
A	0.982	0.711	*
B	1.141	1.123	*
C	0.932	0.926	*
D	*	*	1.172
E	*	*	1.215
F	0.820	0.355	*
G	*	*	1.113
H	1.148	*	*
I	*	1.068	*

ERIC

Table A6

Parent Interview
Weights assigned to each District/Program Combination

District	Program		
	Immersion	Early-Exit	Late-Exit
A	1.229	0.945	*
B	1.049	1.065	*
C	0.846	0.813	*
D	*	*	1.172
E	*	*	1.072
F	1.596	0.321	*
G	*	*	0.885
H	1.216	*	*
I	*	1.321	*

Table A7

Student Data Sheet
Weights assigned to each District/Program Combination

District	Program		
	Immersion	Early-Exit	Late-Exit
A	0.847	0.762	*
B	0.924	0.963	*
C	0.959	0.980	*
D	*	*	0.920
E	*	*	1.180
F	0.981	0.533	*
G	*	*	0.931
H	1.433	*	*
I	*	1.412	*

Table A8

District Services
Weights assigned to each District

District	
A	1.280
B	1.806
C	0.835
D	0.404
E	1.804
F	0.470
G	0.792
H	0.944
I	0.666

Table A9

Project Administrator
Weights assigned to each District

District	
A	0.925
B	1.304
C	0.603
D	0.584
E	2.606
F	0.339
G	1.144
H	1.364
I	0.962

Table A10

School Site Fact Sheet
Weights assigned to each District

District	
A	1.308
B	1.845
C	0.854
D	2.069
E	1.537
F	0.801
G	0.675
H	0.438
I	0.851

Each school received equal weight within a district whether it had one program or two.

Table A11

Principal Interview
Weights assigned to each District

District	
A	1.281
B	1.807
C	0.836
D	2.023
E	1.290
F	0.784
G	0.661
H	0.430
I	1.667

Because the data were duplicated for schools with two programs, the data for each school with two programs were each given one-half the weight listed above. This resulted in all schools within a district receiving equal weight.

ERIC

APPENDIX B

Classroom Language Use Patterns In Study Year 1

ERIC

DIRECTORY OF TABLES

ERIC

APPENDIX B

CLASSROOM LANGUAGE USE PATTERNS IN STUDY YEAR 1

Do the language use patterns observed in study year 1 match the program models?

Yes, they do. Table B1 shows dramatic differences in the proportion of English and Spanish used between the three programs consistent with the study's operational definitions for each. Kindergarten teachers in the immersion strategy program almost always use English (92%) and occasionally use Spanish (7%). Early-exit kindergarten teachers use English approximately two-thirds of the time (64%) and Spanish about one-third of the time (36%). In marked contrast, teachers in late-exit programs only use English about one-third (34%) but use Spanish about two-thirds (66%) of the time. First-grade immersion teachers use English almost exclusively (99.7%) and Spanish seldom (0.3%). First-grade early-exit teachers use less English than kindergarten early-exit teachers. Thus, first-grade early-exit teachers appear to use English slightly more than half of the time (55.4%) and Spanish somewhat less (44%). In sum, teachers in the three instructional programs differ in the use of English and Spanish. These differences are consistent with the study's operational definition for each program.

ERIC

Table B1

Proportion of Teacher Utterances by Language, Program, and Grade
Study Year 1

Grade		Program: Immersion Strategy	Early-Exit	Late-Exit
K	% English:	92.5	63.6	33.0
	% Spanish:	7.4	36.2	65.8
	% Mixed:	0.1	0.2	1.2
1	% English:	99.7	55.4	-
	% Spanish:	0.3	44.1	-
	% Mixed:	0.0	0.6	-
3	% English:	-	-	48.3
	% Spanish:	-	-	51.4
	% Mixed:	-	-	0.3

AGUIRRE INTERNATIONAL

FINAL REPORT: LONGITUDINAL STUDY OF STRUCTURED ENGLISH IMMERSION STRATEGY, EARLY-EXIT AND LATE-EXIT TRANSITIONAL BILINGUAL EDUCATION PROGRAMS FOR LANGUAGE-MINORITY CHILDREN

VOLUME II

Contract No.
300-87-0156

J. David Ramirez
David J. Pasta
Sandra D. Yuen
David K. Billings
Dena R. Ramey

February 1991

Submitted to:

U.S. Department of Education
Washington, D.C.

Aguirre International
411 Borel Avenue, Suite 402
San Mateo, CA 94402
(415) 349-1842

ERIC

FINAL REPORT:
LONGITUDINAL STUDY OF IMMERSION STRATEGY, EARLY-EXIT AND LATE-EXIT TRANSITIONAL BILINGUAL EDUCATION PROGRAMS FOR LANGUAGE-MINORITY CHILDREN

VOLUME II

Contract No.
300-87-0156

J. David Ramirez
David J. Pasta
Sandra D. Yuen
David K. Billings
Dena R. Ramey

February 1991

Submitted to:

U.S. Department of Education
Washington, D.C.

Aguirre International
411 Borel Ave., Suite 402
San Mateo, CA 94402
(415) 349-1842

The research reported herein was performed pursuant to Contract No. 300-87-0156 with the U.S. Department of Education. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official U.S. Department of Education position or policy.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

ERIC

ACKNOWLEDGEMENTS

Aguirre International gratefully acknowledges the assistance of Dr. Ricky Takai, Project Officer of the Office of Planning, Budget, and Evaluation, U.S. Department of Education, and the members of the study's Technical Advisory Panel who supplied constructive guidance at key stages during the study. Dr. John Willett, Professor, Harvard University, Dr. Anthony Bryk, University of Chicago, and especially Dr. Stephen Raudenbush, Professor, Michigan State University, were invaluable in assisting us with implementation of the Hierarchical Linear Model. Drs. Henry Braun, Educational Testing Service, Wendy Yen, CTB-McGraw Hill, and Michael Smith, Dean, School of Education, Stanford University, were extremely helpful in the formulation of the final analytic plan. Most importantly, we are highly appreciative of the analytic and statistical support provided and contributions made by DMA Corporation through the services of Mr. David J. Pasta and Dr. David K. Billings, particularly in effecting the HLM and TAMP analyses.

We would also like to thank the district superintendents, program directors, principals, administrative personnel, and especially the teachers who were so cooperative in arranging schedules, completing questionnaires, permitting interviews, and access to records. The quality of any evaluation rests on the quality of the data. Local school personnel were outstanding in their efforts and support to help us realize the study's goals. This cooperation is greatly appreciated.

Members of the Technical Advisory Panel

Dr. Wallace Lambert
Professor, McGill University
Montreal, Canada

Dr. Eunice Lear
Project Director
San Diego Unified School District
San Diego, CA

Dr. Robert Linn
Professor, University of Illinois
Champaign, IL

Dr. Eva Pena-Hughes
Project Director
McAllen Independent School District
McAllen, TX

Dr. Merrill Swain
Director, Modern Language Center
Ontario Institute for Studies in Education
Toronto, Ontario, Canada

Dr. Gloria Zamora
Project Director, I.D.R.A.
San Antonio, TX

We would like to acknowledge the herculean efforts of project staff members without whom this challenging study could not have been effected:

Field staff

Rolando Castillo	Site Coordinator
Yolando De Duran	Site Coordinator
Carmen Dinos	Site Coordinator
Rosemary Fonseca	Site Coordinator
Teresa Hermosillo	Assistant Site Coordinator
Mary Jo Karam	Assistant Site Coordinator
Antonia Lacen	Assistant Site Coordinator
Carol Lay	Site Coordinator
Jorge Nakamoto	Site Coordinator
Vivian Rodriguez	Site Coordinator
Yolanda Suris	Site Coordinator
Estella Tristan	Assistant Site Coordinator
Martha Vallejo	Assistant Site Coordinator

A special note of appreciation is extended to Ms. Anne F. Abrams, Research Associate, for her skill and unwavering support in the preparation of this volume.

DIRECTORY OF TABLES

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

Page

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

ERIC

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF TABLES
(Continued)

DIRECTORY OF FIGURES

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

Page

ERIC

DIRECTORY OF FIGURES
(Continued)

ERIC

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

ERIC

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

DIRECTORY OF FIGURES
(Continued)

I. INTRODUCTION AND BACKGROUND

Introduction

As noted in "Volume I: Longitudinal Study of Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children," the primary study objective is to assess the relative effectiveness of structured English immersion strategy, early-exit, and late-exit transitional bilingual education programs. Two research questions underlie this objective. The first research question is pedagogical: Is an English-only instructional program as effective as a primary language developmental program in helping limited-English-proficient students acquire English language and math skills? The second research question is a public policy question: Are English immersion strategy and late-exit transitional bilingual instructional programs as effective in meeting the needs of limited-English-proficient students as those programs typically funded (i.e., early-exit)? Answers to these questions will help educators better understand how children acquire language and how this learning can be facilitated through formal instruction. This information will also be helpful to policy makers as they try to determine how best to use limited resources to satisfy the needs of a significant rapidly growing sector of our student population that is limited-English-proficient.

The study objective is realized by determining the extent to which the educational needs of limited-English-proficient students are met through an instructional program that provides all formal instruction in English and/or one in which formal instruction is provided in the child's primary language. Comparing these two instructional approaches is important in that, despite the positive outcomes of Canadian immersion programs for language-majority children, there currently is insufficient empirical evidence to argue for or against the adaptation of the Canadian immersion approach with minority-language children. For this study, one model of an English-only instructional program (i.e., structured English immersion strategy) and two types of a primary language program were examined. The two primary language programs differ in the amount and duration of instruction in the primary language, one using a limited amount of Spanish for a short time period (i.e, early-exit) and the other a substantial amount of Spanish for an extended time period (i.e., late-exit). A comparison of the structured English immersion strategy and the early-exit programs directly addresses both the pedagogic and the research questions. By including information on the academic growth of students in late-exit instructional programs, we are better able to examine the evidence for the underlying hypothesis of a developmental primary language program, i.e., that a strong primary language developmental instructional program facilitates the acquisition of skills in English language and reading as well as the content areas such as mathematics. To address the two primary research questions, ideally one would look for a number of schools wherein one could find each of the three program models in operation. Such schools would have allowed us to consider all three programs in addressing the pedagogic and public policy questions. This was the intent at the initial planning of the study. Pointedly, the following study questions were to be addressed:

Original Study Questions

- What is the relative effectiveness of structured English immersion strategy, early-exit, and late-exit transitional bilingual education programs in meeting the learning needs of limited-English proficient students with respect to English language proficiency and non-language skills (i.e., mathematics)?

- How does the growth of students in structured English immersion strategy, early-exit, and late-exit programs compare to the norming population?

- What generalizations can be made about the education of limited-English-proficient students?

ERIC

As discussed in more detail later, the study did not allow for a true experiment wherein one would establish all three programs in a school and randomly select instructors and students. The parameters for the study were such that a national search was made to identify instructional programs that existed in schools that matched the characteristics of the three instructional programs as defined for this study. No school was found in which all three programs were operated. Schools with programs that resembled those defined for this study came either from a district that had an immersion strategy and an early-exit program or one that had only a late-exit program. A completely crossed design allowing for effective control of district, school, and classroom characteristics was not possible. As a result, the original study questions were modified to reflect the availability of the three instructional programs of interest:

Final Study Questions

- What is the relative effectiveness of a structured English immersion strategy and an early-exit transitional bilingual education program in meeting the learning needs of limited-English-proficient students with respect to English language proficiency and non-language academic skills (i.e., mathematics)?

- How did each of the three late-exit programs compare in the patterns of student growth across districts, schools, and cohorts within the late-exit program model?

- How does the growth of students in structured English immersion strategy, early-exit, and late-exit programs compare to the norming population?

- What generalizations can be made about the education of limited-English-proficient students?

Drawing from data collected from four years of the study, Volumes I and II of this report provide information on the homogeneity of each instructional program across classrooms within each instructional treatment, the differences and similarities in classrooms among treatments, and the comparability of student/community characteristics within and among instructional treatments. Student achievement data are presented as trends. The public policy and research context for this study as well as the characteristics of each instructional program are detailed in Volume I. Volume II documents the success with which each program meets the needs of limited-English-proficient students. These data provide valuable information to policy makers and practitioners about alternative approaches to bilingual education and the requirements for the successful implementation of each.[1]

Other Project Reports

The Longitudinal Study of Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children began in FY 1983-84 and ended in FY 1990-91. Year one of the project realized four major tasks: (a) finalizing the study design (Ramirez et al., 1984); (b) developing data collection instruments; (c) preparing literature reviews (Ramirez, Schinke-Llano, & Bloom, 1984; Schinke-Llano & Ramirez, 1984); and (d) selecting study sites (Ramirez, Wolfson, & Morales, 1985). Year two of the project was the first year of a four-year data collection effort producing information on the students in the study and their instructional programs (Ramirez et al., 1984). Years three and four of the study resulted in the Second and Third Year Reports wherein data describing students and each instructional program collected were summarized and reported separately by year (Ramirez et al., 1987, 1988). These yearly reports allowed us to describe the incremental

[1]Please see "Study Design of the Longitudinal Study of Immersion Programs for Language-Minority Children" for a description of the policy context for this study. This study, the history of federal bilingual education policy, and a summary of recent research and evaluation studies of instructional services to limited-English-proficient students are described in this report.

changes in the instructional programs of the target students as they moved into the higher grades.

Purpose of the Report

Whereas Volume I of the final report was concerned with the issue of confirming instructional treatment, "Volume II: Final Report of the Longitudinal Study of Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children" describes the achievement of Spanish-speaking language-minority children participating in these instructional programs over time. This volume presents information on the relative effectiveness of structured English immersion strategy, early-exit, and late-exit transitional bilingual education programs implemented at the study sites. The analyses take into account programmatic and contextual information that identify differences and similarities among these three programs.

The success of any instructional program that is designed to address the needs of students who are performing below the achievement of the general population must ultimately be evaluated in terms of how well the special instruction helped the students perform relative to the norming population. (The norming population is defined as the universe of individuals to whom another population is compared. In this instance, as defined by the CTBS test, the norming population is all kindergarten through sixth grade students in the United States.) To this end, Volume II also relates the growth in academic achievement of language-minority children participating in each of these programs to the growth of the general student population.

To provide a context for this study, Volume I began with a brief overview of the history of federal bilingual education policy and a summary of recent research and evaluation studies of instructional services designed for limited-English-proficient students. To focus the discussion of student achievement, the remainder of this chapter outlines the theoretical definition for each instructional program (i.e., Nominal

Program) and presents data confirming the implementation of each instructional program (i.e., Operational Program). The chapter concludes with an overview of the organization of the succeeding chapters.

Defining the Instructional Treatments

Major criticisms levied against prior studies of bilingual programs include the failure to document the instructional treatment, and combining programs that represent distinct instructional models. Understandably, failure to consider the unique characteristics of instructional treatment yields results that are not meaningful or conclusive. To assure ourselves on this important issue, a specific instructional model was defined for each instructional program. A census was taken looking for those programs that most resembled these definitions. This search revealed a diversity of implementation for each nominal program label (Ramirez et al., 1985). Given this diversity, it is critical that the reader be continually aware of the three specific treatments defined and selected for this study. This section begins with a theoretical definition (i.e., of nominal program) for each program model, followed by data confirming the actual implementation (i.e., operational program definition).

Theoretical Model (Nominal Program Definition)

As defined in this study, English immersion strategy, early-exit, and late-exit transitional bilingual education programs are differentiated primarily on the basis of the language(s) used in the classroom and the amount each is used. Table 1 details the characteristics of each program model. While these program characteristics are discussed in detail in Volume I, it is appropriate to review them once again, as they are essential for understanding the achievement data.

A draft of the characteristics for each program were developed after an exhaustive review of the literature on language development theory and first and second language research. This draft was reviewed and refined with the assistance of the study's advisory panel. The advisory panel, comprised of researchers, practitioners, and school administrators, reflected expertise in language development theory, research, practice, and public policy. The final list of program characteristics was reviewed and approved by staff from the Office of Bilingual Education and Minority Language Affairs (OBEMLA) and the Office of Planning, Budget and Evaluation (OPBE), United States Department of Education.

While the characteristics of each program were detailed in Volume I, tt would be helpful at this juncture to review these characteristics once again.

Program Characteristics

Four Methods for Teaching English to Limited-English-Speaking Students

Transitional bilingual education, however, is not the only means of teaching English to limited-English-speaking children in American public schools. Rather, four general alternatives have evolved to serve language-minority children: (a) submersion; (b) English as a second language (ESL); (c) transitional bilingual education (TBE); and (d) structured English immersion strategy. These programs differ in five main areas: (a) whether traditional all-English instruction is used; (b) whether special instruction in English as a subject is provided; (c) whether English is taught through the teaching of other subjects; (d) whether non-language subjects such as mathematics are taught in the primary language of the limited-English-proficient students; and (e) whether the primary language of limited-English-proficient students is used (to supplement instruction in English). The following describes the four programs in terms of these characteristics (see Table 1).

ERIC

Table 1

Characteristics of Programs Serving Language-Minority Children in the United States

Characteristics	Submersion	English as a Second Language	Transitional Bilingual Education	Structured English Immersion Strategy
Traditional all-English instruction used	Yes	Yes	No	No
Special instruction in English as a subject is provided	No	Yes	Yes	No
English is taught through the teaching of other subjects	No	No	Yes	Yes
Informal use of LEP student's primary language for clarification	No	No	Yes	Yes
Non-language subjects are taught in LEP student's primary language	No	No	Yes	No

The submersion approach is typical of the majority of services provided to language-minority students (Office of Bilingual Bicultural Education, 1981). In submersion programs, language-minority children are placed in ordinary mainstream classrooms where only English is spoken. No special provisions are made in these mainstream classrooms to help them learn quickly the English skills they need to succeed in school. Instead, they study the regular curriculum and are expected to perform as best they can. To the extent that the mainstream classroom represents an instructional program wherein English and other academic skills are developed, it can be considered as an instructional alternative to transitional bilingual education (i.e., one in which no special instructional adjustments are made). This alternative has been described as "sink or swim." Placement in submersion classes usually results when there are insufficient numbers of limited-English-proficient students in the same grade and school to "trigger" a bilingual classroom. Many of these students receive English as a second language instruction.

In typical English as a second language (ESL) programs, language-minority students spend most of their day in a submersion classroom, but receive some extra instruction in English. This special instruction is developed specifically to teach English as a second language. For subjects other than English, the language-minority students study the school's standard curriculum in English-only classrooms. In teaching ESL, the teacher may or may not use the primary language of the limited-English-proficient students. The California State Department of Education conducted a recent survey which determined that after submersion programs, ESL is the next most widely used instructional method for teaching language-minority students (Office of Bilingual Bicultural Education, 1981).

In transitional bilingual education (TBE) programs, language-minority students study subject matter in their primary language until they have learned enough English to succeed in English-only mainstream classrooms. Children in TBE programs generally learn to read first in their non-English home language and then in English. ESL often is used as a supplement to reduce the time needed to learn English. TBE programs are similar to submersion and ESL in that English usually is taught as a separate subject, but differ in that other content areas, including reading, are taught in the child's non-English home language, at least in principle.

The structured English immersion strategy program (SEISP) is proposed as an alternative to ESL and TBE programs. It is based on the results of Canadian French immersion programs for language-majority (i.e., English) speakers. While similar to ESL and some TBE programs in many ways, the SEISP also differs from them substantially. All instruction is in English. Rather than teaching English strictly as a subject, however, the SEISP endeavors to teach English through the various content areas. Prior knowledge of English is not assumed. Instead, teachers in SEISPs carefully tailor their English to a level the limited-English-proficient students can understand. For example, a lesson about a particular science concept also would teach the development and use of specific English language skills (such as vocabulary). Content thus becomes the medium for teaching language. Such teaching differs from transitional bilingual programs in that SEISPs present the subject matter exclusively in English, while TBE programs teach content in the students' primary language until they have learned English.

The SEISP teacher is bilingual and speaks the students' non-English home language. The student may use their primary language among themselves and to address the teacher. Generally, however, the teacher speaks to the students in English, using the home language only occasionally to provide or clarify instructions. Understanding the children's home language, the teacher can determine whether a child's difficulty with a given task stems from a problem with the language or with the content itself. As the child's primary language is not used formally for instruction, this type of instruction would be best described as a "sheltered English" program, rather than as a bilingual program.

To ensure a clear understanding of structured English immersion strategy, early-exit, and late-exit programs, the following section identifies those characteristics that the three programs share and those on which they differ.

Characteristics Shared by All Three Programs

o <u>Teachers must have specialized training in language development</u>.

From Table 2 we can see that the program models require that the teacher be bilingual and have specialized language development training (i.e., certified bilingual and/or language development specialist teaching credential). This recognizes that as teacher training programs for the standard teacher credential provide limited or no training in how children develop language skills (speaking, reading, and writing) in their first language, teachers are ill prepared to help second language learners acquire these skills efficiently and successfully. (These comments do not apply to teachers who have obtained specialist credentials, i.e., reading specialist or language specialist, and who only work with native English speakers. They refer to teachers who receive the basic coursework requirements for the general elementary teaching credential and teachers with specialist credential_ who have not been trained on how to address the needs of language-minority students.)

Exacerbating the problem of limited language training that teachers receive is the fragmentation of training that typically occurs within most training institutions. Training in the content areas, such as in history, health, science, etc., tend to focus only on those concepts that are relevant to each area. Rarely are teachers taught how the content areas can be used to develop and reinforce basic language skills.

Thus the general elementary teacher, lacking a language development framework, not only lacks an understanding of how the native English speaking student learns to read, write, and speak, but also how these skills interrelate, and how they can be developed through the various content areas. The typical teacher possessing a general teaching credential would have even less understanding of how to address these needs among children who do not have English as their first language. Yet these are precisely the teaching skills needed by teachers of second language learners.

ERIC

Table 2

Characteristics of Program Models

Characteristics	Immersion Strategy	Early-Exit	Late-Exit
Teachers must have specialized training in language development.	X	X	X
Teachers must be bilingual.	X	X	X
Students are free to use their primary language (i.e., L1) among themselves and with the teacher.	X	X	X
Content areas are used to teach L2 (i.e., English).	X	X	X
L2 is used to teach content.	X	X	X
Substantial English language component.	X	X	X
Teachers use both L1 and L2 for instruction.		X	X
Teachers have native or near native language skills in L1 and L2.		X	X
Children are mainstreamed as soon as they are reclassified as being fluent-English-proficient (i.e., approximately 2-3 years in program).	X	X	
Teachers' use of L1 is informal (limited in amount and frequency).	X		
All instruction is done in English.	X		
Limited primary language development component.		X	
Limited instruction in content areas in L1.		X	
L1 language arts skills are developed before or at the same time as L2 language arts skills are developed.		X	
Substantial primary language arts component.			X
L1 language arts skills are developed first, prior to introducing L2.			X
Substantial use of L1 for instruction in amount (> 50% of instructional day) and duration (through grade 6).			X
Regardless of when children are reclassified, they are not mainstreamed until after grade 6.			X
L1 language arts skills are developed before L2 language arts skills are introduced.			X

o Teachers must be bilingual.

It is clear that early-exit and late-exit teachers must be bilingual in that they must spend part of their day providing instruction in their students' primary language. However, it might not be as clear why English immersion strategy teachers who provide instruction only in English also must be bilingual. While the immersion strategy model requires that the language of instruction primarily be English, it also requires that the second language learner be allowed to respond in his first language until the point where his second language skills are sufficiently developed to begin speaking it. This last feature of the model is included in recognition of the time required to develop receptive language prior to being able to speak. Consequently, English immersion strategy teachers must be bilingual to allow them to understand and respond appropriately, albeit in English, to students who answer questions in Spanish.

o Students are free to use their primary language (i.e., L1) among themselves and with the teacher.

This is a requirement of each instructional program for two reasons. First, it recognizes the developmental stages one goes through prior to producing language as outlined above. Second, it is a direct way for teachers to validate the child's background, and allows for many opportunities to help the child develop and maintain positive self-esteem. Both reasons are critical for successful learning.

Given the implicit and explicit social status that teachers possess by virtue of their position, their receptivity and response to children when children use their primary language with one another or when they communicate with the teacher conveys recognition and approval for such communication. In short, it validates the legitimacy of children's efforts and it provides an opportunity for children to demonstrate what they have learned without having to struggle with second language issues. For example, the learning atmosphere is very supportive of a limited-English-proficient student when, in response to the question "What is two

ERIC

times two?." the student is free to say "Dos," and the teacher reinforces the student with "Wonderful, that is well done." If, however, the teacher were to respond with "No, say it in English," the message communicated to the student is one of error. And if the student is in the initial stages of learning English and has not yet mastered the word "Two," the learning situation would result in communicating failure to the student.

- <u>Content areas are used to teach L2 (i.e., English)</u>.

This requirement is inherent in each model as it recognizes that language learning is enhanced when there is a language development component to each content lesson. By maintaining a clear understanding and focus of how different instructional activities support the development of language skills, teachers can better select appropriate learning activities and reinforce language learning all day rather than just during the one period of English as a second language (ESL). Every interaction with the student can become a language learning experience.

- <u>L2 is used to teach content</u>.

Using L2 to teach content is required in each model in recognition that language learning is facilitated when language is used purposefully. In this instance, L2 learning increases when the student attempts to use it to convey meaning about a particular topic. Using L2 as the language of instruction provides a reason for the student to learn the content; if he/she does, the student will be able to communicate in class.

- <u>Substantial English language component</u>.

On the average across grades, at least forty percent or more of the total instruction is provided in English. English instruction includes development of oral, written, and reading skills as well as for use in instruction of content areas such as, but not limited to, science, mathematics, and/or social studies.

Characteristics Common Only to Early-Exit and Late-Exit Models.

- Teachers use both L1 and L2 for instruction.
- Teachers have native or near native language skills in L1 and L2.

Characteristics Common Only to English Immersion Strategy and Early-Exit Models.

- Children are mainstreamed as soon as they are reclassified as being fluent-English-proficient (i.e., approximately 2-3 years in program).

Characteristics Unique to the English Immersion Strategy Model.

- Teachers' use of L1 is informal (limited in amount and frequency).
- All instruction is done in English.

Characteristics Unique to the Early-Exit Model.

- Limited primary language development component.
- Limited instruction in content areas in L1.
- L1 language arts skills are developed before or at the same time as L2 language arts skills are developed.

Characteristics Unique to the Late-Exit Model.

- Substantial primary language arts component.
- L1 language arts skills are developed first, prior to introducing L2.
- Substantial use of L1 for instruction in amount (> 50% of instructional day) and duration (through grade 6).

- *Regardless of when children are reclassified, they are not mainstreamed until after grade 6.*

- *L1 language arts skills are developed before L2 language arts skills are introduced.*

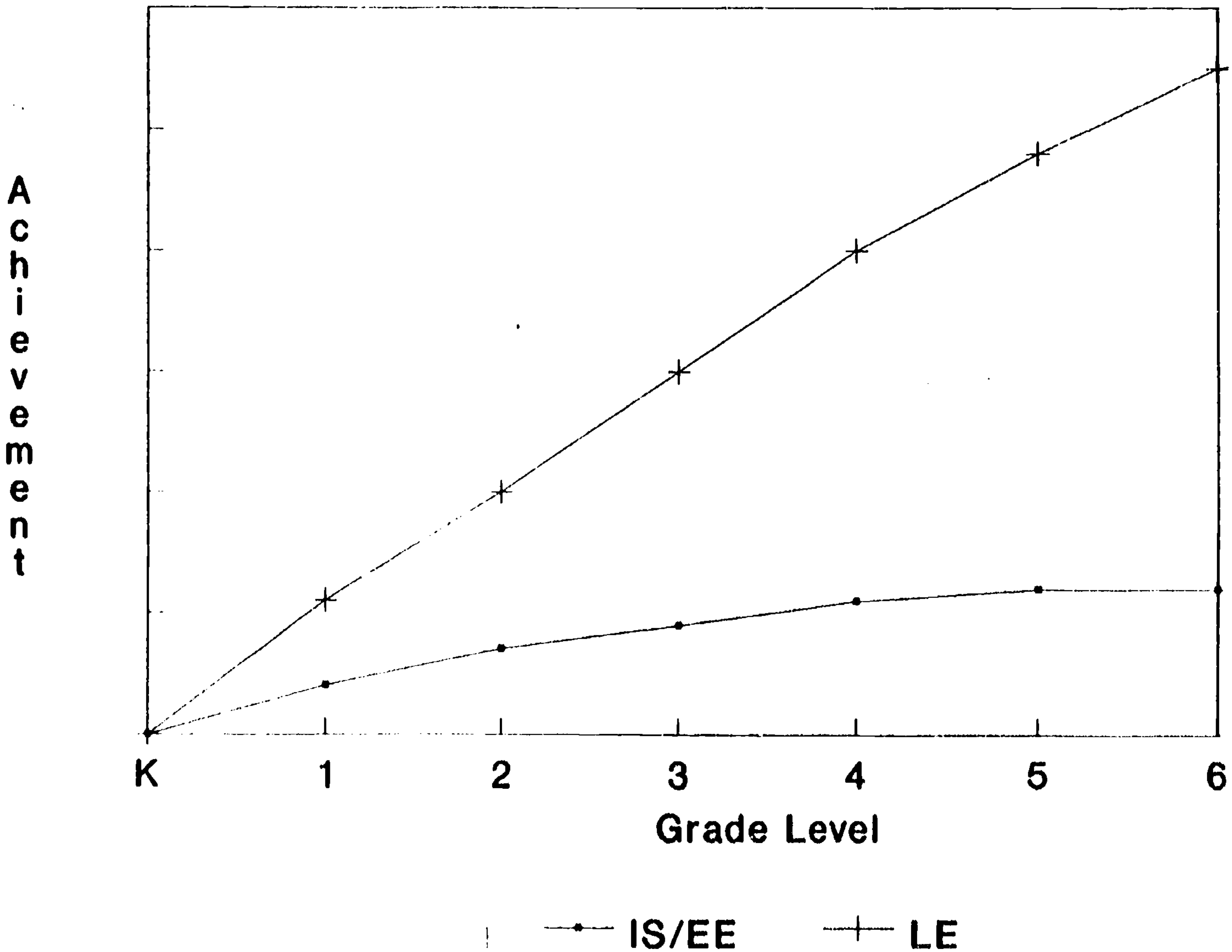
In sum, the major differences between the three models are in the amount and length of time that L1 is used for instruction and when a student is to be mainstreamed. Each of these models assumes certain patterns of academic growth. The following section outlines the achievement that would be predicted from each of these three instructional models.

Hypothesized Effectiveness of Each Instructional Program

The primary study task is to determine whether the achievement of immersion strategy students is comparable to that of early-exit students, and to describe the achievement of students in alternative late-exit programs. To this end, student achievement was assessed in English for mathematics, language arts (i.e., grammar/mechanics), and reading. As the instructional model of each program is different, differential growth rates in each of these areas were predicted for each model. Figure 1 through Figure 4 are sketches of the predicted qualitative differences in expected growth for each instructional model. As such, they are not to be evaluated as quantitative predictions, but merely as illustrations of hypothesized student growth.

Figure 1

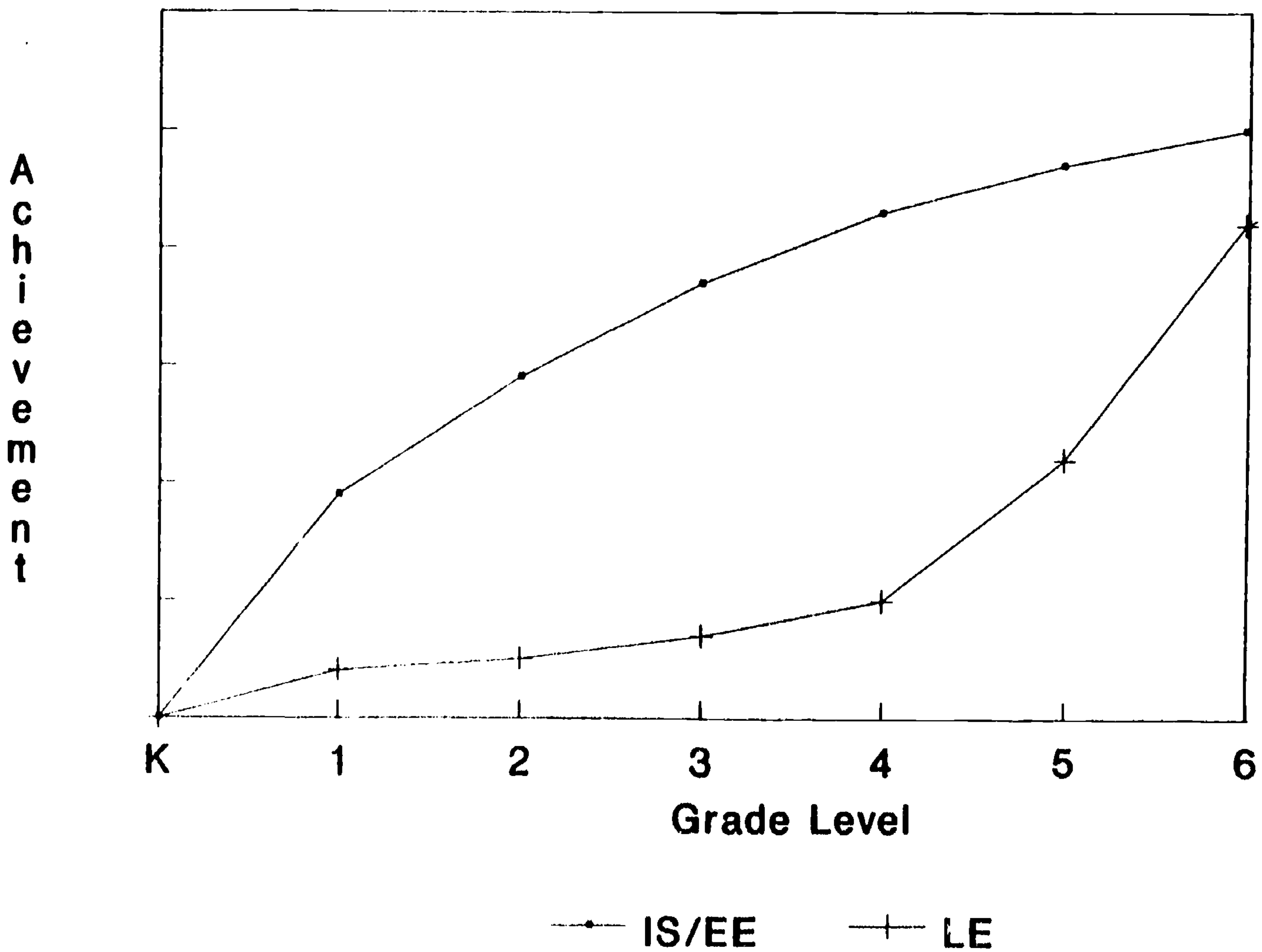
Hypothesized Primary Language Proficiency



IS = Immersion Strategy   EE = Early-Exit   LE = Late-Exit

Figure 2

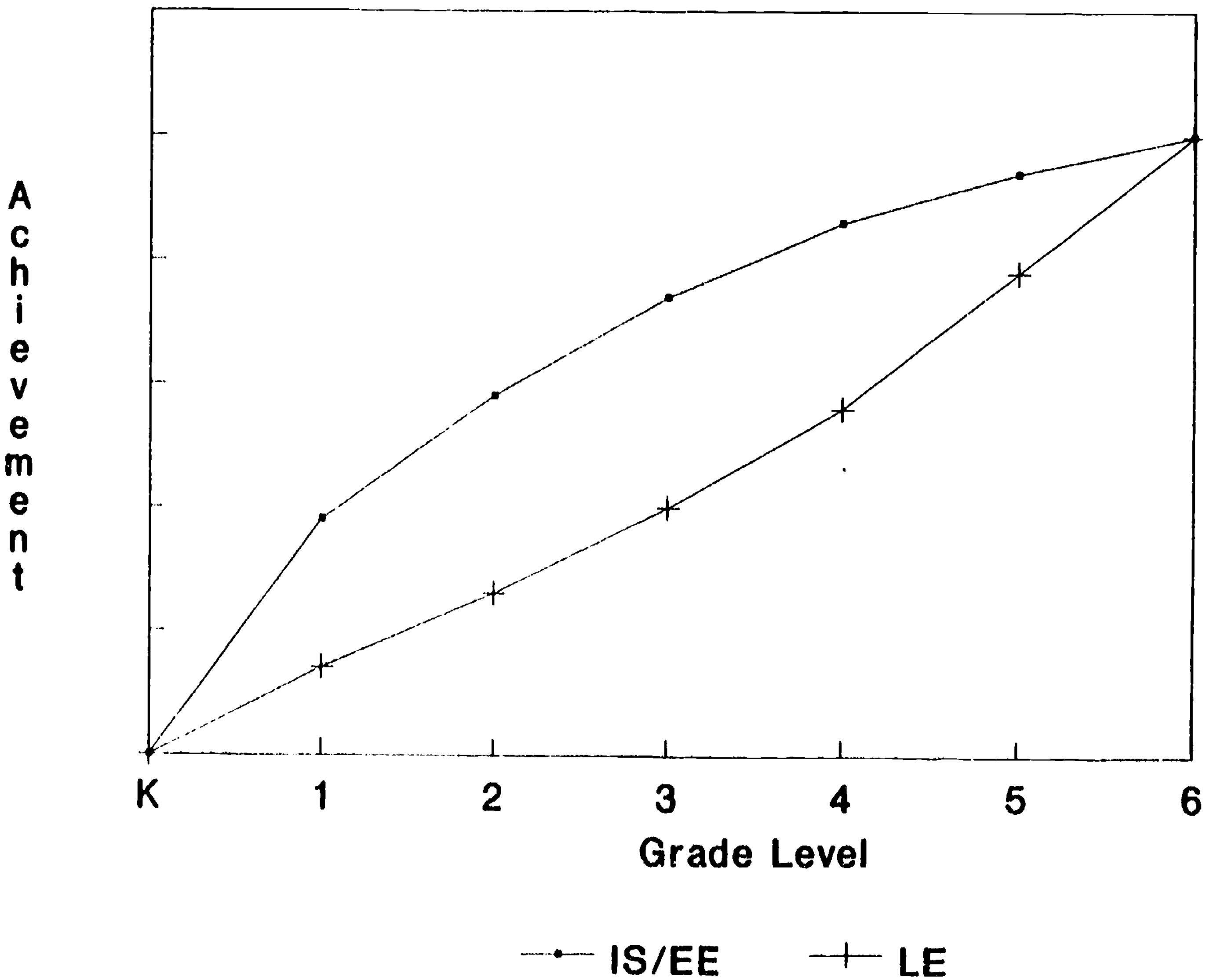
Hypothesized English Reading and Writing Proficiency



IS = Immersion Strategy EE = Early-Exit LE = Late-Exit

Figure 3

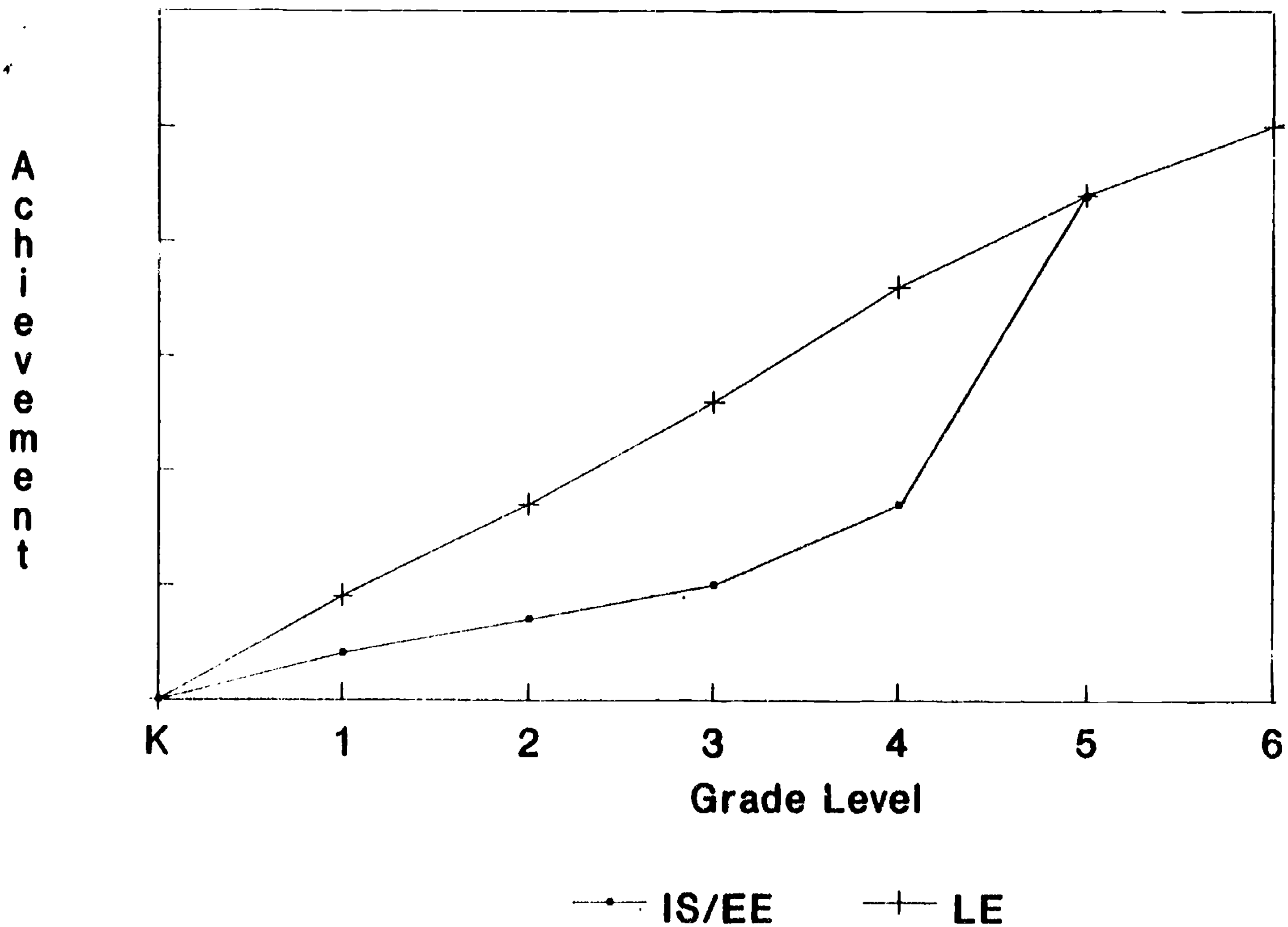
Hypothesized Oral English Proficiency



IS = Immersion Strategy EE = Early-Exit LE = Late-Exit

Figure 4

Hypothesized Mathematics and Social Studies Achievement



IS = Immersion Strategy EE = Early-Exit LE = Late-Exit

As the proportion and duration of Spanish language instruction are limited in early-exit classrooms, it was predicted that the instruction in an early-exit program would so closely resemble that of an immersion strategy program that the achievement of early-exit students would be comparable to the achievement of immersion strategy students. It also was predicted that as limited-English-proficient students in these two programs would be provided with more English initially, the achievement of immersion strategy and early-exit students in English language arts and reading would be greater than the achievement of late-exit students in the early primary grades. On the other hand, it was predicted that the facilitative effects of substantial primary language development and instruction upon English language proficiency as provided in late-exit programs would begin in grade four and not be evident until grades five and six. Thus, while in the short run (i.e., kindergarten through grade four) it was predicted that immersion strategy and early-exit students would demonstrate greater proficiency in English language arts and reading than late-exit students, late-exit students would catch up and demonstrate comparable skills by grades five and six. Also, the prediction was made that, because students' primary language would be used for instruction in the content areas, late-exit students would out-perform immersion strategy and early-exit students in the content areas such as math in the early primary grades. However, given the models of each program, it was predicted that immersion strategy and early-exit students would be as proficient in these content areas as late-exit students by grades five and six.

Testing for these predictions requires gathering data for students in each program from kindergarten through grade six. Two constraints precluded such data collection. First, the study was limited to gathering data over a four-year period. This did not allow for following a cohort of entering kindergarten students for seven years. To ensure that this grade span was covered, two cohorts of students were needed, one beginning in kindergarten and followed through grade three, and a second cohort (who had also begun the program in kindergarten) followed from third grade through sixth grade. This pattern of data collection was carried out for students in the late-exit program. However, at the time that the study was initiated (1983), fewer than six viable structured English immersion strategy programs existed in the United States. They were so recently implemented (most within two years) that there were insufficient numbers of third grade students to form a second cohort. As the goal for immersion strategy and early-exit programs is to mainstream their students as soon as possible, usually by the end of first or second grade, it was determined that an optimal plan would follow a kindergarten to third grade cohort and a first grade to fourth grade cohort (of students who had been in the program since kindergarten). This would allow these students to be followed in the mainstream classes for one or two years. Thus the study design allows for information on the growth of immersion strategy and early-exit students from kindergarten through fourth grade and late-exit students from kindergarten through sixth grade.

In sum, the predictions outlined above flow from each of the instructional models. One of the first steps in the analyses is to confirm the instructional treatment. That is, to what extent do the classrooms within each program reflect their respective instructional models? The following discussion answers this question.

Actual Implementation (Operational Definition)

At this point it is important to differentiate between two types of analyses: nominal and operational. Nominal program analyses seek to compare programs on the basis of their "program label." In this study it means collapsing data across districts and schools that were collected from programs initially identified as either immersion strategy, early-exit, or late-exit. Operational program analyses would use the data gathered from direct observation of teacher behaviors to determine the appropriate program category for analyses. For example, a program in a given district initially may have been identified and self-described as an early-exit program, but after observation it is determined that no teacher in this program used the child's primary language more than five percent of the time and provided all instruction in English. In this instance, this program would be recategorized as an immersion strategy program. That is, districts, schools, and/or classrooms would be regrouped according to specific features that were observed to be in common. This reassignment based on observations might or might not agree with the original categorization.

The first step in any analysis is to effect the nominal program analysis. That is, given the a priori categorization of type of program, programs in different districts described as adhering to a given instructional model are compared to programs in the same or other districts described as adhering to a different instructional model without concern for within-program variation in either group. If between-group differences are not found, then the next round of analyses would directly address the issue of consistency of program implementation among districts within the same program label. If it were discovered that variation in implementation was quite large, this might warrant subdividing the districts into more coherent program clusters and then using the nominal program analysis as a guide to program differences based on these new clusters. If program differences are discovered at the nominal program level, then the operational analyses would allow identification of those characteristics at the district, school, or classroom level that seem to have the most impact upon the outcome. Once again, the results of the nominal program analyses guide the direction and scope of the operational program analyses. The primary analytic effort as required for this study is essentially a nominal program analysis. Nonetheless, it is extremely helpful to gauge the conclusiveness of the results of the nominal program analysis if there is an understanding of the fidelity of treatment. That is, to what extent do the programs in this study reflect their instructional model and how comparable are they to one another within their respective program labels? The following provides this information.

ERIC

What is the purpose of the operational program analyses?

The nominal program analyses focus on the relative effectiveness of the immersion strategy and early-exit programs and a comparison of student achievement across districts within the late-exit program. In contrast, operational program analyses are designed to evaluate the educational programs students experience without regard to the name of the program. That is, the operational program analyses seek an operational definition of effective programs by identifying predictors of growth rate (and initial status) other than the program label. A complete operational program analysis, designed to establish the most effective educational programs for language-minority students, is beyond the scope of this report. Given the richness of the data collected, such an analysis is an enormous undertaking. The operational program analyses presented here have the limited goal of determining the extent to which the amount of English used in the classroom explains all or most of the identified differences in achievement when tested in English in the nominal programs.

What operational programs were identified?

As noted earlier, the proportion of English used in the classroom by the teacher is the major characteristic differentiating the three instructional models. However, as in the implementation of any model, variance from the "ideal" model occurs as a result of the unique characteristics and resources of the school district in which it is implemented. (See Appendix A for a listing of schools by program within each district.) Consequently, it is important that one obtain a clear understanding of the implementation of each model. To this end, the patterns of English use in each program were examined to identify similarities in the English use patterns among the programs and to determine the degree of homogeneity within each nominal program. Eight different operational programs were identified through their unique patterns of English language use (see Table 3). A standard deviation of twenty percent or greater is used as a conservative benchmark to note significant differences.

Table 3

Average Percentage of English Use for Instruction in Classroom

Pattern	Program		K	1	2	3	4	5	6
1	Immersion Strategy, all	mean+ S.D.++	99.39 ± 2.79	98.84 ± 2.34	99.10 ± 2.56	99.80 ± 3.82	100.00 ± 7.42	* *	* *
2	Early-Exit, Site F	mean S.D.	* *	* *	98.92 ± 8.82	99.92 ± 9.42	99.07 ± 10.32	* *	* *
3	Early-Exit, Sites A, I	mean S.D.	70.26 ± 3.82	70.95 ± 3.03	91.53 ± 3.95	99.04 ± 5.31	100.00 ± 9.84	* *	* *
4	Early-Exit, Site B	mean S.D.	66.07 ± 4.72	64.76 ± 3.98	69.28 ± 5.13	80.81 ± 6.70	95.79 ± 12.05	* *	* *
5	Early-Exit, Site C	mean S.D.	54.44 ± 5.07	56.85 ± 6.28	53.36 ± 7.55	28.19 ± 11.54	* *	* *	* *
6	Late-Exit, Site G	mean S.D.	* *	55.08 ± 6.40	56.50 ± 7.12	71.85 ± 7.76	83.06 ± 7.36	94.98 ± 9.29	96.29 ± 10.15
7	Late-Exit, Site D	mean S.D.	40.05 ± 7.42	56.62 ± 8.42	58.29 ± 9.55	69.49 ± 8.33	56.83 ± 11.30	50.13 ± 12.33	67.28 ± 12.97
8	Late-Exit, Site E	mean S.D.	16.84 ± 3.64	25.09 ± 4.26	12.89 ± 4.92	39.05 ± 4.13	55.02 ± 5.05	60.54 ± 6.20	77.06 ± 11.54

\+ Estimated population mean (least squares mean) (see SAS Institute Inc., 1985, for description of General Linear Models [GLM] procedure).

++ Pooled estimated standard deviation from ANOVA.

* No data available at this grade level.

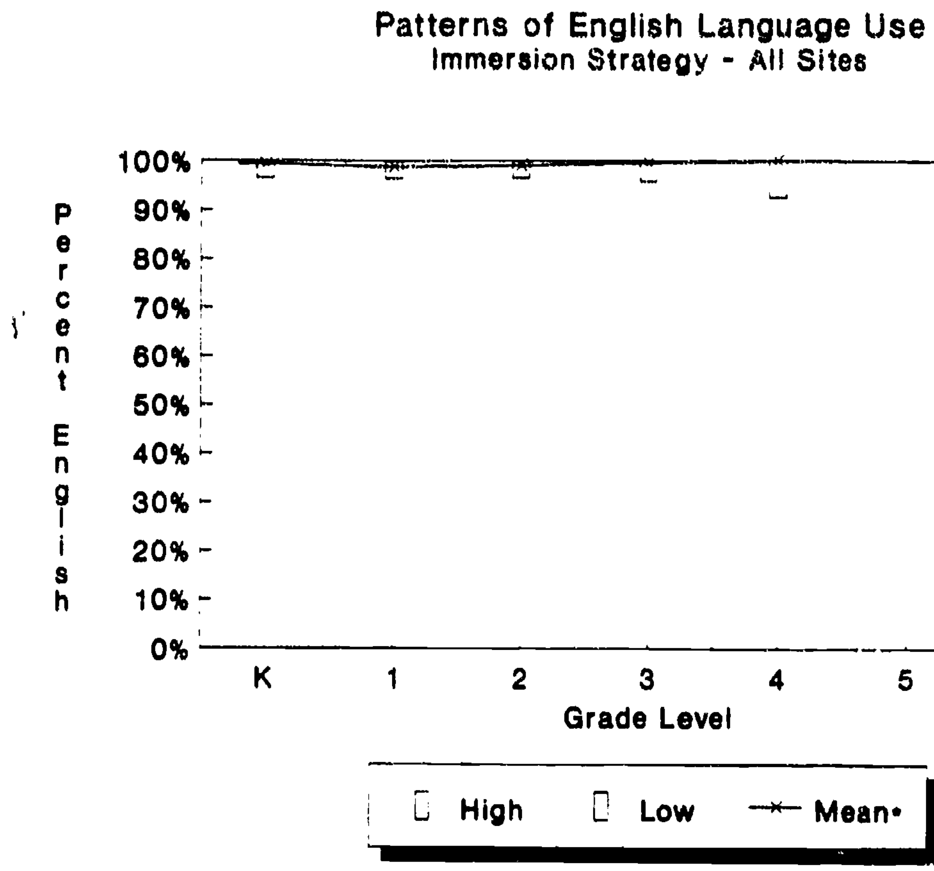
Pattern 1: A striking outcome of the pattern identification is the finding that, of the three instructional programs, only teachers in the immersion strategy program sites are homogeneous in their use of English. That is, all immersion strategy teachers across sites are consistent in their near exclusive use of English from kindergarten through grade four. Consistently across grades, English is used on the average from 98.84% to 100% of the time. The lack of variability of English language use among immersion strategy teachers is evident in that the standard deviations range from no more than 2.34% to 7.42% (see Figure 5).

Pattern 2: Early-exit site F classrooms are practically indistinguishable from immersion strategy program classrooms in that teachers at this site also use English almost exclusively, averaging from 98.92% to 99.92% of the time. However, the standard deviation is greater than that found in immersion strategy classrooms, as it ranges from 8.82% to 10.32% (see Figure 6).

Pattern 3: Two early-exit sites (sites A and I) have similar patterns of English use. Teachers use English in these sites slightly less than three-fourths of the time in kindergarten and first grade. There is a dramatic increase in the use of English in grade two as it increases to 91.53%. In grades three and four, English is used almost exclusively. Unlike Pattern 2, these sites are fairly homogeneous in that the standard deviations are very small, ranging from no more than 3.03% to 9.84% (see Figure 7).

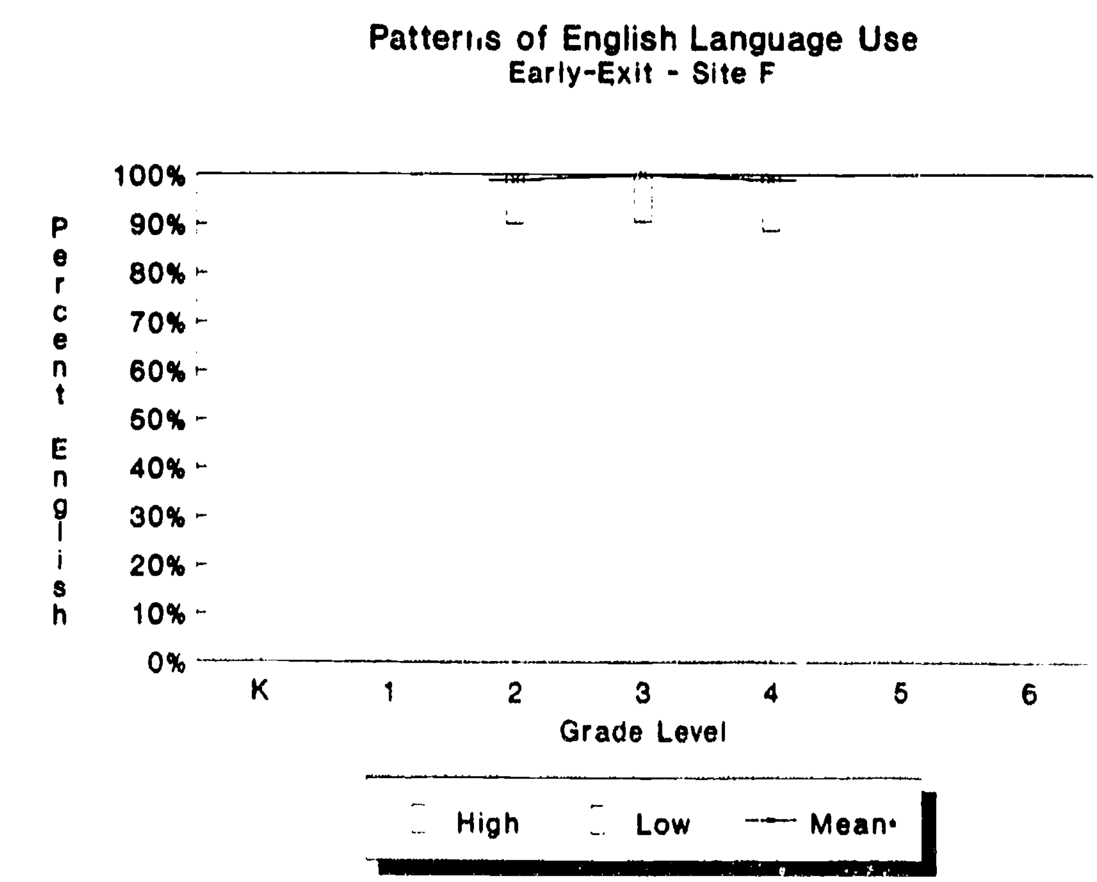
Pattern 4: Teachers in early-exit site B show a more modest use of English and a slightly smaller increase in its use after grade two. English is used roughly two-thirds of the time in grades kindergarten through two, then increases to 80.81% in grade three. Not until grade four is English used almost exclusively (95.79%). Once again, the variance within this program by grade is not very great, as the standard deviations range from a low of 3.98% to a high of 12.05% (see Figure 8).

Figure 5



• Mean • Least squares mean

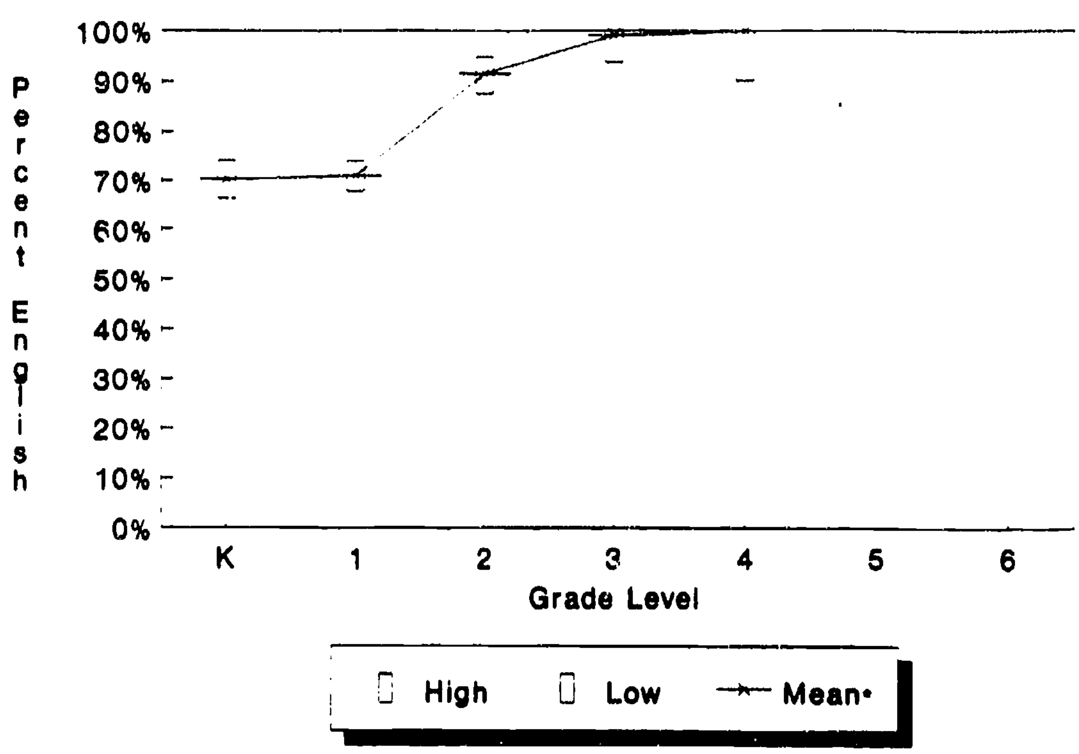
Figure 6



• Mean • Least squares mean

Figure 7

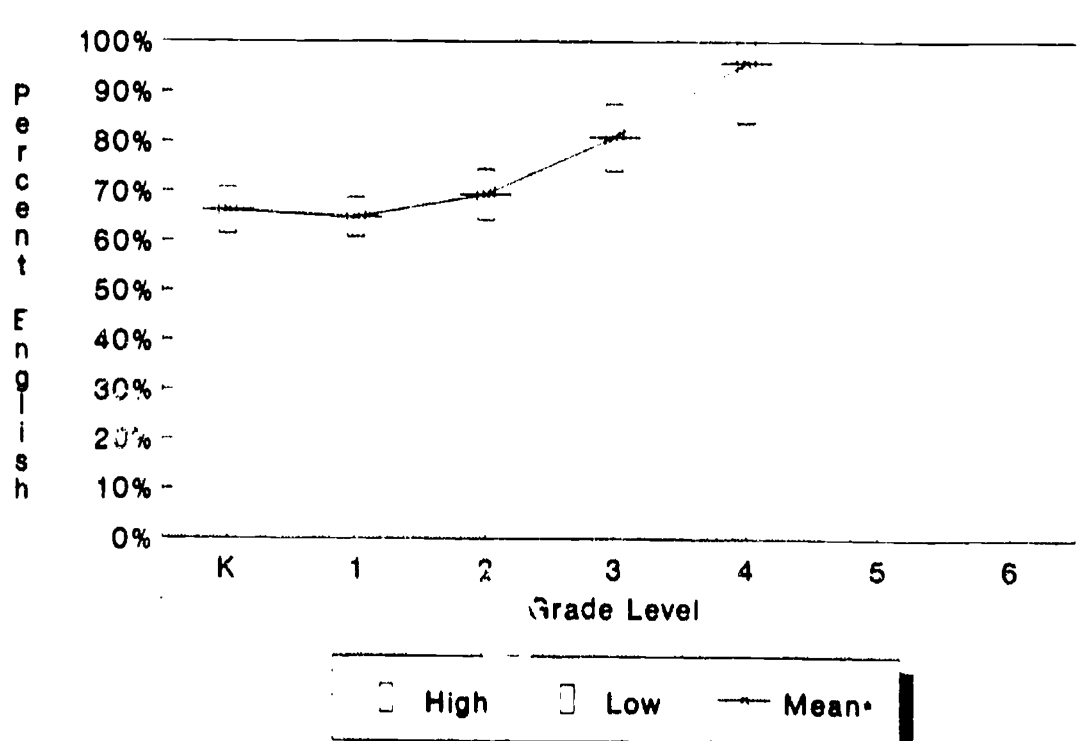
Patterns of English Language Use
Early-Exit - Sites A and I



• Mean • Least squares mean

Figure 8

Patterns of English Language Use
Early-Exit - Site B

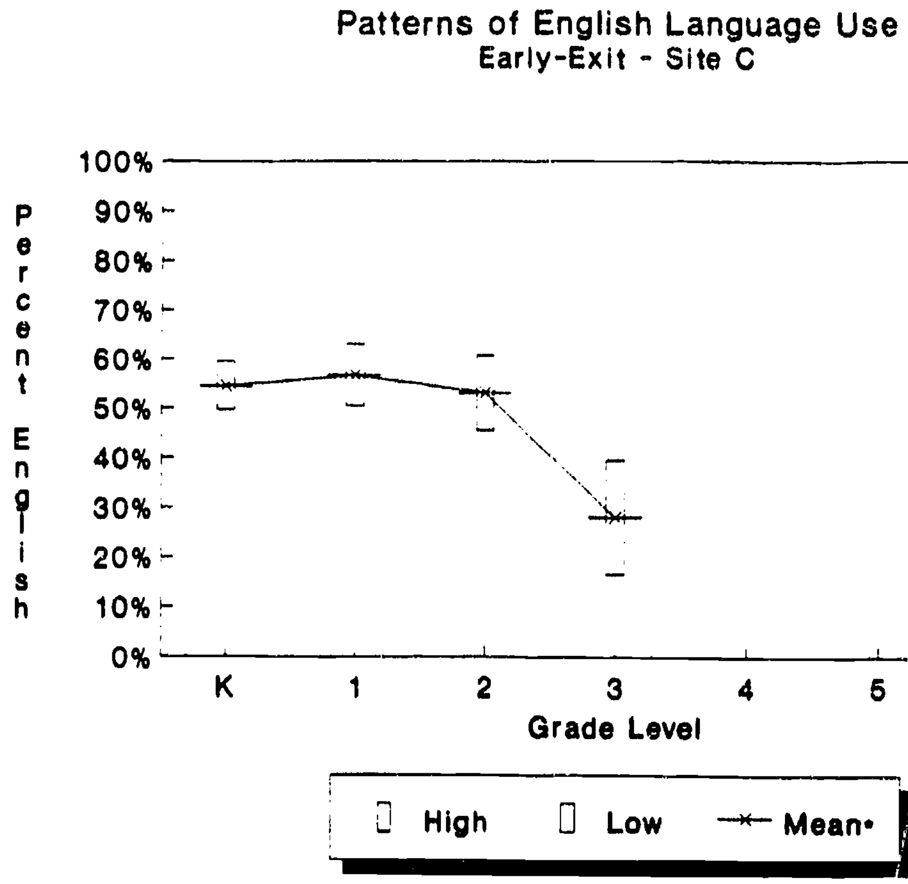


• Mean • Least squares mean

Pattern 5: Early exit site C in some ways appears to resemble a late-exit site more than a site with an early-exit program, in that English is used only slightly more than half of the time in kindergarten through second grade. Unlike any other site, the use of English declines substantially in grade three to 28.19%. This uneven use of English and the reduction in English in grade three reflect a marked degree of inconsistency in this program site not only in terms of its nominal program label, but in an unevenness in curriculum which, in turn, might have an adverse effect upon student learning (see Figure 9).

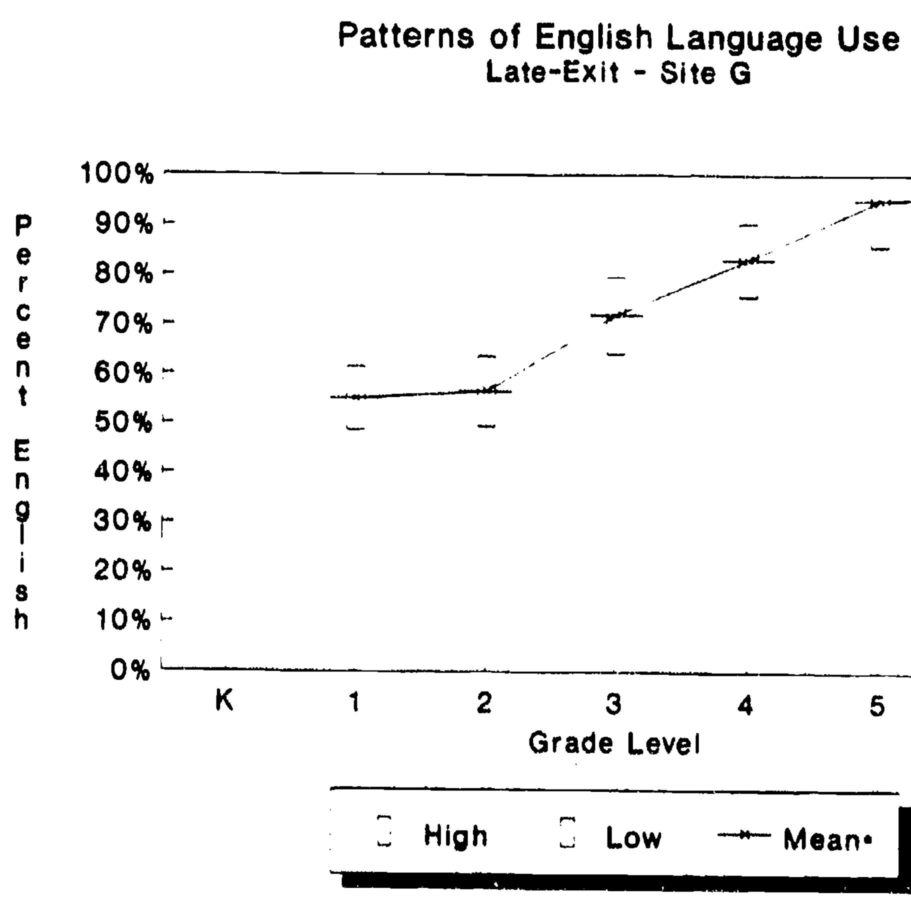
Pattern 6: While late-exit site G teachers are true to the nominal program label through the second grade in that English is used less than 60% of the time, by grade three the site resembles that of an early-exit program site, as English is used 71.85% of the time. English is used four-fifths of the time in grade four. By grades five and six, teachers at this site resemble teachers in immersion strategy and early-exit program sites in that English is the principal language of instruction. A look at the standard deviations reveals a bit of variance within classrooms at this site, as they range from 6.4% to 10.15%. In sum, late-exit site G is more similar to an early-exit program site than to a late-exit program site (see Figure 10).

Figure 9



* Mean = Least squares mean

Figure 10



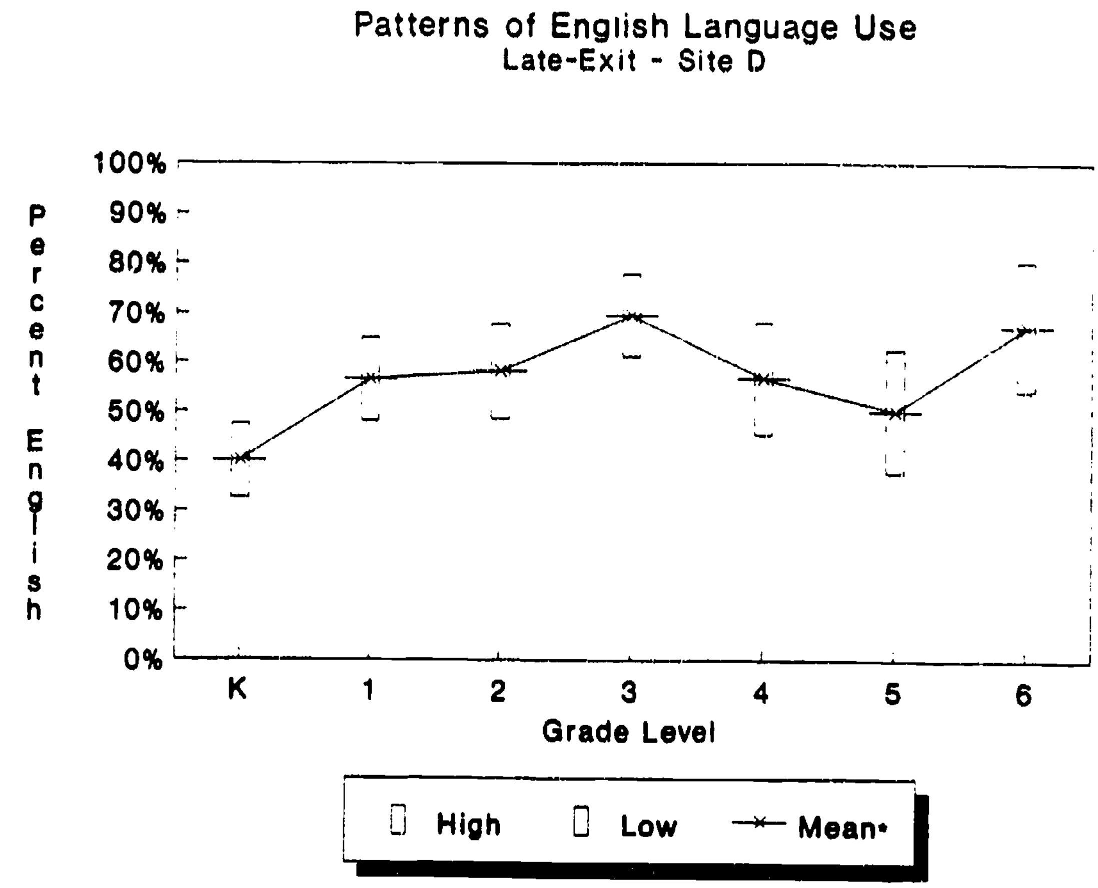
* Mean = Least squares mean

Pattern 7: With two exceptions, the program at site D is consistent with its late-exit nominal program label across grades. English is used about two-fifths of the time in kindergarten, then slowly increases to slightly more than one-half in first and second grades. Contrary to the late-exit model, English is used nearly two-thirds of the time in grade three, only to return to the late-exit model once again in grades four (56.83%) and five (50.13%). In grade six, English again is used two-thirds of the time. While the pattern of English use is somewhat uneven, on the average it is not far out of range. Of all of the patterns, this site exhibits the greatest amount of variation (i.e., heterogeneity of variance) within its classrooms, as the standard deviations range from 7.42% to 12.97%. Moreover, the instructional model calls for use of English not to exceed 60%. In the two instances that English use exceeded this amount, the differences were marginal. In sum, site D appears to be a late-exit site (see Figure 11).

Pattern 8: Site E is the late-exit site that most consistently reflects its nominal program designation. English is used sparingly in kindergarten (16.84%), then increases to about one-fourth in grade one. However, contrary to expectations, the use of English drastically declines to 12.89% in grade two, only to increase to about two-fifths in grade three. By grade four, English is used slightly more than one-half of the time, and up to the 60% limit established in the definition of the late-exit model in grade five. However, in grade six, English is used nearly three-fourths of the time, exceeding the limits of English use established for this program model. Nonetheless, site E most closely resembles its late-exit label (see Figure 12).

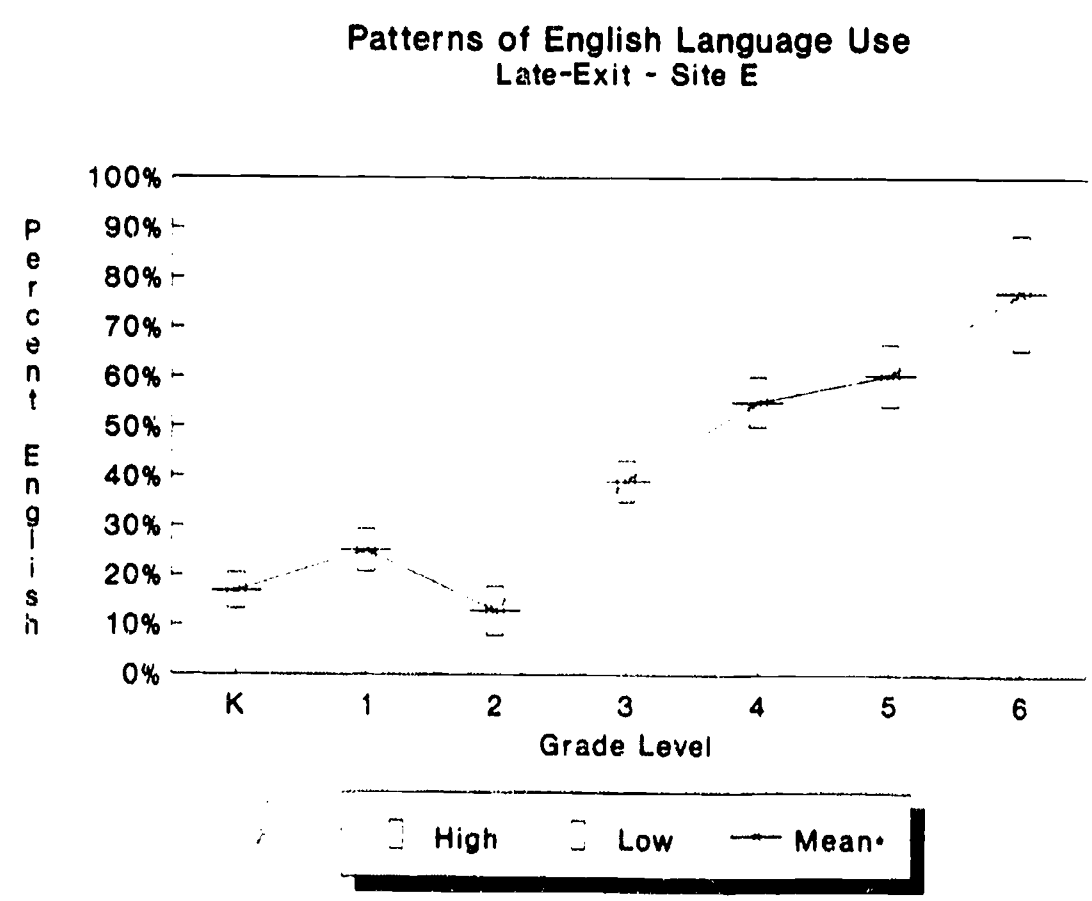
ERIC

Figure 11



• Mean • Least squares mean

Figure 12



• Mean • Least squares mean

ERIC

Across all eight patterns, the variation in amount of English language use tends to increase with grade level, with the greatest amount of variation occurring in grades four and above.

In sum, the immersion strategy program sites exhibit the greatest fidelity to their nominal program label and the least amount of variation within program. There is a great deal of diversity within both the early-exit and late-exit programs. The program at one early-exit site more closely resembles an immersion strategy program than an early-exit program in its almost complete reliance upon English across all grades. Two early-exit sites begin with substantial use of English and increase to almost total English use by grade two. Another early-exit site begins with large proportions of English, but exhibits a smaller increase in the use of English. The final early-exit site more closely resembles a late-exit program than an early-exit program in that English instruction is used less than 60% of the time across grades. The program at one late-exit site more closely resembles an early-exit program in that the use of English exceeds the nominal program specifications in grade three and above. Another late-exit site, while resembling its nominal program label overall, exceeds the 60% limit on English use, although the difference is minimal. The remaining late-exit site faithfully implements the instructional model through grade five. In grade six, however, English is used more than the limit established for this model. The differences in patterns of English use between each of these patterns are notable. It is recommended that the relative effectiveness of each of these observed patterns in meeting the needs of the limited-English-proficient students they serve be examined. Unfortunately, this goes beyond the resources of this study, but is an effort that should be undertaken in the future.

When limited-English-proficient students should be mainstreamed is another central issue differentiating the instructional programs. According to their respective models, limited-English-proficient students in the immersion strategy and early-exit programs are to be mainstreamed upon reclassification to fluent English proficiency. This is predicted to occur by the end of first or second grade. Late-exit students reclassified as fluent-English proficient are not to be mainstreamed until after the sixth grade.

As reported in Volume I, reclassified late-exit students, consistent with the instructional model, were not mainstreamed until after grade six. (By the end of the fifth grade, approximately four-fifths of the students had been reclassified as fluent-English-proficient.)

Contrary to the respective models, after four years in their respective programs, only about two-thirds of the immersion strategy and three-fourths of the early-exit students had been reclassified. Thus, students in the early-exit program were reclassified at a higher rate than those in the immersion strategy program. Even more surprisingly, although teachers in both programs were committed to early mainstreaming, after four years in their respective programs, only one-fourth of the immersion strategy students and less than one-fifth of the early-exit students had been mainstreamed. In effect, not all immersion strategy or early-exit students who were reclassified were mainstreamed. This presumably reflected the judgments of their teachers and principals that although these students had met local reclassification criteria, their skills were still not sufficiently strong to function in a mainstream English-only classroom. Clearly, as these students had been reclassified, there was no financial incentive to maintain these students in these programs other than to ensure that their learning needs were met. This factor should not be forgotten as the achievement results are reviewed, for it suggests that limited-English-proficient students require a minimum of five or more years of special instruction in either immersion strategy or early-exit bilingual programs. That is, the achievement results for immersion strategy and early-exit students reflect this additional training beyond that proposed by the instructional models.

This chapter provided an overview of the study objectives and research questions, as well as information on the definition of each treatment (i.e., English immersion strategy, early-exit, and late-exit transitional bilingual education programs). Information was provided on the theoretical background (Nominal Definition) and the actual implementation (Operational Definition) of each instructional treatment. As the purpose of Volume II is to address the student growth in three specific instructional programs, the predictions of expected student outcomes for each treatment were discussed.

At this juncture, we know that the three programs in this study do in fact differ substantially from one another, and that these differences, with one exception, are consistent with their respective instructional models. While there was a diversity of patterns in the amount of English used among early-exit sites in grades kindergarten through second grade, they all had moved to almost exclusive use of English by third grade. With two exceptions, all roughly fit under the nominal program definition.

The one late-exit site that more closely resembled an early-exit program is more of a concern. The pattern of English use is so different from the nominal program label and from the other two late-exit sites after grade three that it might affect the achievement pattern for the late-exit students if students from this site were combined with those of the other two late-exit sites. The effect of these differences in patterns of English among the early-exit sites as well as among the late-exit sites would best be addressed in future analyses of operational programs.

Organization of the Report

The chapters that follow address the issue of the relative effectiveness of the three instructional programs. A study with a complex design and the typical difficulties of longitudinal data collection in field settings pose challenging analytic issues. Some of these issues were expected, while others arose as the analyses proceeded. The result has been a series of complex analyses. Chapter II enumerates each analytic step as well as a number of the analytic issues that arose and how they were addressed.

Chapter III addresses the question of the relative effectiveness of immersion strategy and early-exit programs for schools that had both programs. These analyses represent a completely crossed design at the school level.

Chapter IV also addresses the question of the relative effectiveness of immersion strategy and early-exit programs, but these analyses use data from schools that had only an immersion strategy or an early-exit program. These analyses represent a comparison of two groups of schools.

Chapter V addresses the growth of students in each of the three late-exit districts within this instructional program. These analyses also relate achievement for the group of students in the early primary grades to the achievement for another group of students in the later primary grades.

Ultimately, the success of any instructional program for language-minority children is the degree to which students receiving these support services are able to achieve at a level commensurate with the general population. To this end, Chapter VI relates the achievement of students in each of the eight districts and three instructional programs to each other and to the achievement of the general student population.

Chapter VII presents conclusions and implications.

ERIC

II. OVERVIEW OF THE ACHIEVEMENT ANALYSES

Introduction

The intent of this chapter is to provide an overview of the steps taken to analyze the student achievement data. The chapter is divided into two parts. Part one enumerates each of the major sections, and corresponding tasks, of the data analyses. Part two is a discussion of a range of key analytic issues that were addressed.

What was the purpose of the analyses of achievement tests?

The analysis of data as complex as that generated by this study is necessarily incomplete, but the richness of the collected information does permit us to offer substantial answers to our main questions. In this volume we focus on one and only one question: what differences in student achievement, as measured in English, were found among the three programs? The analyses that follow focus on the programs according to their names: immersion strategy (IS), early-exit (EE), or late-exit (LE).

Data analyses comparing instructional programs can be done at both the nominal program level and the operational program level. A nominal analysis is a comparison of instructional programs according to their respective instructional "labels." These categories are those which programs are given or to which programs are assigned on the basis of surface program characteristics, such as funding source. In contrast, an operational analysis is a comparison of instructional programs by categories to which they have been assigned based on observed differences in the instructional program, such as the proportion of instruction provided in English, whether content instruction is used to teach second language skills, etc. A nominal program analysis addresses the question of whether there are any differences in outcomes between program categories. An operational program analysis addresses two questions: (1) What are the different categories into which the instructional programs can be grouped based on observed differences in program characteristics?, and (2) Are there differences in outcomes between these different types of programs?

An operational program analysis is generally preceded by a nominal program analysis. A nominal program analysis is important because it provides valuable information on the relative effectiveness of different programs, taking into account the range of implementations of each label. A given instructional model may have a single definition, but when implemented a range of interpretations and accommodations will follow, providing a richer and more complex definition. It is this diversity in implementation of a given model that is taken into account in the nominal program analysis. Thus, if a nominal program analysis results in significant differences between two types of programs, one would be confident that despite the variation in instructional treatment within the programs, the program difference is sufficiently large to warrant distinguishing between the programs.

Another reason a nominal analysis is an important first step before an operational analysis is that an outcome of the nominal analysis is the development of an analytic plan. This analytic plan provides the framework for the operational program analysis. The contribution of the operational program analysis is that by disregarding the nominal classification, it categorizes program implementations by key instructional characteristics, thereby minimizing the within-category variation. This approach allows one to identify and focus on those specific features of the instructional approach which will maximize student learning. It allows one to find out how each program operates and how any differences in program effects can be explained by differences in program implementation.

The primary task of this study is to effect a nominal analysis of the three instructional programs. The analyses that follow address this critical analytic step. For some analyses, differences in student growth across districts are related to differences in the program implementations across those districts. However, a complete operational program analysis is beyond the scope of this study.

Analytic Framework

The achievement analyses were divided into four sections:

1. A comparison of the immersion strategy and early-exit programs in the schools that had both programs (see Chapter III).

2. A comparison of the immersion strategy and early-exit programs in the schools that had only one of the programs (see Chapter IV).

3. A comparison of the three implementations of the late-exit program (see Chapter V).

4. Comparisons among all districts and all programs through the use of special graphical techniques (see Chapter VI).

In the IS/EE analysis of the two-program schools, it was possible to estimate the program effect while controlling for the effects of school. In the IS/EE analysis of the one-program schools, it was possible to control for district. Although a comparison of the LE program with the IS and EE programs was not possible because no district with a LE program had either of the other two programs, a graphical technique called TAMP (Trajectory Analysis of Matched Percentiles) was used to evaluate the relative gains of groups of students in various districts and programs.

A K-1 analysis was completed for the two-program IS/EE schools, the one-program IS/EE schools, and the three LE districts. The K-1 analyses consisted of a series of analyses of covariance, using kindergarten test scores as a covariate, which permitted a comparison of the programs at spring of first grade.

A growth curve analysis of grades 1-3 was completed for the two-program IS/EE schools, the one-program IS/EE schools, and the three LE districts. Additional late-exit analyses included a growth curve analysis for grades 3-6, and a growth curve analysis of grades 1-6. The growth curve analyses consisted of individual growth curve modeling.

Each analysis was subjected to a series of sensitivity tests. The stability of the major findings was evaluated by adding a number of covariates, by varying the model, by varying the subjects included in the analysis, and by varying the analytic method. In addition, separate analyses were performed to evaluate the effectiveness of the kindergarten pretest to predict the growth of achievement in grades 1-3.

Table 4 summarizes the major analyses, and Table 5 outlines the specific analytic procedures for these major analyses and identifies the chapters where each is reported.

Table 4

Major Analyses

	Immersion Strategy and Early-Exit		
	Two-Program	One-Program	Late-Exit
K-1 Analysis	All Students Pretest Only	All Students Pretest Only	All Students Pretest Only
1-3 Analysis	All Students Pretest Only	All Students Pretest Only	All Students
3-6 Analysis			All Students
1-6 Analysis			All Students

Table 5

Analytic Procedures for the Major Analyses

	K-1 ANALYSES Analysis of Covariance (ANCOVA) and Trajectory Analysis by Matched Percentiles (TAMP)	1-3 (1-6) ANALYSES Hierarchical Linear Models of individual growth curves (HLM) and Trajectory Analysis by Matched Percentiles (TAMP)
Chapter III. IS/EE Two-Program Schools Analyses: Comparison of IS students and EE students in schools with both programs.	ANCOVA: Compare programs at spring first grade, controlling for school effects.	HLM: Compare programs at spring first grade and on growth rate spring first grade to spring third grade, controlling for school effects.
Chapter IV. IS/EE One-Program Schools Analyses: Comparison of students in schools having only the IS program with students in schools having only the EE program.	ANCOVA: Compare programs at spring first grade, controlling for district effects but treating schools as a random effect. Control for selection and attrition effects using "propensity scores."	HLM: Compare programs at spring first grade and on growth rate spring first grade to spring third grade (1) treating schools as a fixed effect, and (2) treating schools as a random effect while controlling for district effects. For both sets of analyses, control for selection and attrition effects using "propensity scores."
Chapter V. LE Districts Analyses: Comparisons among the districts (D, E, and G) and cohorts (K-3 and 3-6).	ANCOVA: Compare districts at spring first grade, treating schools as a random effect (K-3 cohort only).	HLM: Compare districts and cohorts at spring third grade and on growth rate (spring first grade to spring third grade for the K-3 cohort; spring third grade to spring sixth grade for the 3-6 cohort), treating schools as a fixed effect.
Chapter VI. Trajectory Analysis by Matched Percentiles (TAMP): Comparisons among the three programs, the thirteen district-program combinations (five districts with both IS and EE programs and three districts with LE programs), and the national norms.	TAMP: Compare the students in the national norming sample, the students in each of the three programs, and each of the thirteen district-program combinations on growth from fall kindergarten to spring kindergarten; on growth from spring kindergarten to spring first grade; and on growth from fall kindergarten to spring first grade.	TAMP: (1) Compare the students in the national norming sample, the students in each of the three programs, and each of the thirteen district-program combinations on growth from spring first grade to spring third grade. (2) Compare the students in the national norming sample and the three LE districts on growth from spring third grade to spring sixth grade (3-6 cohort only). (3) Compare the students in the national norming samples and the three LE districts on growth from spring first grade to spring sixth grade (across the K-3 and 3-6 cohorts).

BEST COPY AVAILABLE

The following is a brief but more detailed description of the analyses in Chapters III, IV, and V. The remainder of Chapter II consists of a detailed description of the design of the analyses, the methods used, the test instruments, and the students involved in the analyses.

Comparison of Immersion Strategy and Early-Exit Programs: Two-Program Schools

* K-1 Analysis of Covariance
    * All Students Eligible for K-1 Analysis
        * Program
        * Program and School
        * Program and School with Key Covariates
        * Program and School with All Covariates

    * Only Students Eligible for K-1 Analysis and with Pretest Scores
        * Pretest Not Used as Covariate
            * Program
            * Program and School
            * Program and School with Key Covariates
            * Program and School with All Covariates

        * Pretest Used as Covariate
            * Program
            * Program and School
            * Program and School with Key Covariates
            * Program and School with All Covariates

    * Only Students Eligible for K-1 Analysis and 1-3 Analysis
        * Program
        * Program and School
        * Program and School with Key Covariates
        * Program and School with All Covariates

* 1-3 Growth Curve Analysis
        * All Students Eligible for 1-3 Analysis
        * Only Students Eligible for 1-3 Analysis and with Pretest Scores
            * Pretest Not Used as Covariate
            * Pretest Used as Covariate

Comparison of Immersion Strategy and Early-Exit Programs: One-Program Schools

* K-1 Analysis of Covariance
    * All Students Eligible for K-1 Analysis
        * Program
        * Program and Propensity
        * Program with Key Covariates
        * Program with All Covariates
        * Program and District
        * Program, District and Propensity
        * Program and District with Key Covariates
        * Program and District with All Covariates

    * Only Students Eligible for K-1 Analysis and with Pretest Scores
        * Pretest Not Used as Covariate
            * Program
            * Program and Propensity
            * Program with Key Covariates
            * Program with All Covariates
            * Program and District
            * Program, District and Propensity
            * Program and District with Key Covariates
            * Program and District with All Covariates

        * Pretest Used as Covariate
            * Program
            * Program and Propensity
            * Program with Key Covariates
            * Program with All Covariates
            * Program and District
            * Program, District and Propensity
            * Program and District with Key Covariates
            * Program and District with All Covariates

* 1-3 Growth Curve Analysis
    * All Students Eligible for 1-3 Analysis
        * Basic Two-Level Model
        * Model with School Curvature
        * Models with and without District Main Effects

    * Only Students Eligible for 1-3 Analysis and with Pretest Scores
        * Pretest Not Used as Covariate
        * Pretest Used as a Predictor of Achievement
        * Pretest as a Predictor of Propensity and Initial Status

ERIC

Analysis of Late-Exit Program

* K-1 Analysis of Covariance
    * All Students Eligible for K-1 Analysis
        * No Covariates
        * Key Covariates
        * All Covariates
    * Only Students Eligible for K-1 Analysis and with Pretest Scores
        * Pretest Not Used as Covariate
        * Pretest Used as Covariate
* 1-3 Growth Curve Analysis
* 3-6 Growth Curve Analysis
* 1-6 Growth Curve Analysis

Analytic Issues

How does the study design affect the nominal program analyses?

The analyses that can be performed to evaluate the nominal program are severely limited by the study design, which in turn was severely limited by the paucity of IS and LE programs. Despite an extensive national search, only a handful of immersion strategy and late-exit programs were identified. Essentially all of the districts with IS or LE programs agreed to participate in the study. However, no district with a late-exit program had any other program. This fact, together with the limited number of districts with late-exit programs, makes comparison of the LE program with the IS and EE programs extremely difficult. That is, one cannot control for district or school level differences. Although operational program analyses might be able to overcome some of the difficulties, in nominal program analyses any differences between the achievement of students in LE districts and those in the IS/EE districts cannot be unequivocally attributed to the program: the differences might be due to some other difference between the LE districts and the IS/EE districts.

It can be argued that the inability to control for district differences between the LE program and either the IS or EE program implies that no such comparisons should be made. In the presence of dramatic differences between the growth rate in achievement of LE students and IS/EE students relative to the norming population, however, some comparisons can and should be made. In addition, some insights into the effect of LE programs can be obtained by comparing the three LE districts among themselves. The absence of any 3-6 cohorts of IS or EE students means that it is not possible to compare the 3-6 achievement of LE students with the 3-6 achievement of students in other bilingual programs. These students can only be compared across the three LE districts, with the K-3 cohorts from the LE districts, and with the national norms.

Although it is possible to control for district when comparing the IS and EE programs, it is generally not possible to control for school. All five of the study districts with immersion strategy programs also had early-exit programs in the district (or, in one case, in a neighboring district). This permits nominal program analyses to compare IS and EE programs while controlling for district. However, we found much variability among the schools within each district. This school variability makes it desirable to control for school as well as district. Comparisons of the IS and EE programs in the schools that implement only one program cannot control for school, and most schools implement only one program. In contrast to the extremely limited number of LE districts, however, there were enough one-program IS or EE schools to permit relating program differences to the school-to-school variability.

There were four study schools that had implemented both IS and EE programs. In these four schools, a comparison of the programs can be made while controlling directly for school (and therefore district) effects. This is by far the most satisfactory part of the design for performing nominal program analyses.

Why is it important to control for district and school in nominal program analyses?

It is important to control for both district and school in nominal program analyses not only because those institutions have direct educational effects but also because districts and schools indirectly represent communities. Each district draws its limited-English-proficient (LEP) students from a relatively homogeneous population, while different districts may be drawing upon very different LEP populations. This is perhaps most striking with regard to the country of origin of the Spanish-speaking LEP populations in the study districts.

Districts differed markedly in the percentage of students (9% to 31%) and their parents (40% to 91%) born outside the United States, as well as by their country of origin (see Table 6). In five districts (sites A, B, C, F, and G), the majority of students and their parents who were born outside of the United States came from Mexico. In contrast, most students in three districts and their parents who were born outside the U.S. came from Puerto Rico or South America. Non-U.S.-born students and parents in site D typically came from Cuba, whereas in sites E and H/I they were born in South America or Puerto Rico.

The socioeconomic status (SES) of the families served also varied considerably from one district to another. This is an important consideration, for the SES level of the school suggests the degree of need for support services and the availability of resources to the school. Considerable school-to-school variation in SES was found within districts, as well as the expected variation between districts. As the SES of the students in a school is consistently and strongly associated with achievement in the school, any such variation must be controlled to permit sensible program comparisons. However, a measure of the general SES level for the school was not always available.

Table 6

Percentage of Students and Their Parents
Born Outside of the United States

District	Student	Parents	Country of Origin
A	9%	57%	Mexico
	<1%	3%	Cuba
B	9%	91%	Mexico
	0	1%	South America
C	17%	91%	Mexico
	0	1%	South America
D	24%	72%	Cuba
	7%	10%	South America
E	18%	55%	Puerto Rico
	10%	3%	South America
F	7%	40%	Mexico
G	15%	86%	Mexico
	1%	4%	Puerto Rico
H	15%	35%	Puerto Rico
	13%	33%	South America
	3%	0	Cuba
I	11%	28%	South America
	9%	21%	Puerto Rico
	2%	16%	Cuba
	1%	0	Mexico

In some studies, a random sample of students from a school is obtained, thus permitting the SES of the sampled students to be used as an estimate of the overall school SES. In this study, however, only LEP students were sampled. While for some schools and even entire districts nearly all students are LEP, in other schools and districts varying proportions of the students are LEP. Furthermore, systematic differences between the LEP and non-LEP students in a school are to be expected. Therefore, in instances where school-level SES information was not available for a given school site, the SES of the LEP students is

ERIC

generally not a good measure of the SES of all the students in the school.[2]

Second language learners are able to increase their acquisition of the new language if there are native speakers available (assuming that there are appropriate opportunities for interactions between them). Thus the relative mix of language-minority students and native English speakers could be an important source of support for learning English. Table 7 details the great variability in the mix of students in classrooms by English language proficiency. In some districts, the classrooms are almost entirely comprised of limited-English-proficient (LEP) students, with few or no students who are fluent in English (fluent-English-proficient, or FEP, and English-only, or EO), such as in districts F, H, and I. In other districts, such as B and C, while the majority of the students are LEP (>50%), about one-fourth of the students are either FEP or EO.

A major requirement for each program is that teachers be bilingual. Teachers' understanding of the primary language of their students allows them to determine the success of a lesson and to select appropriate learning tasks and teaching strategies. While districts are fairly consistent in the number of teachers who speak English fairly well, these teachers differ markedly in their proficiency in their students' primary language. Whereas in district E the range of teachers' proficiency in Spanish is exceptionally high, other districts, such as B and C, have few teachers who are sufficiently proficient to teach in Spanish (see Table 8).

---

[2]Some school-level SES information was collected on the School Site Fact Sheet. These measures of SES were relatively crude estimates, however, with questionable reliability. Even this limited information was missing for a substantial number of the study schools.

Table 7

Ranges of Percentage of LEP[1], FEP[2], and EO[3] Students, by District and Program

District		Immersion Strategy %	Early-Exit %	Late-Exit %
A	LEP	37.0 - 100	33.7 - 87.8	*
	FEP	0 - 61.8	1.8 - 57.5	*
	EO	0 - 1.2	8.8 - 21.1	*
B	LEP	53.6 - 62.4	53.8 - 72.1	*
	FEP	9.6 - 12.8	7.5 - 18.2	*
	EO	28.0 - 34.4	20.3 - 28.0	*
C	LEP	54.1 - 65.9	57.6 - 89.6	*
	FEP	4.1 - 8.3	6.3 - 14.5	*
	EO	26.3 - 39.4	0 - 28.7	*
D	LEP	*	*	0 - 89.9
	FEP	*	*	10.1 - 97.7
	EO	*	*	0 - 4.2
E	LEP	*	*	24.4 - 89.6
	FEP	*	*	10.4 - 75.6
	EO	*	*	0
F	LEP	94.7 - 98.4	69.4 - 92.8	*
	FEP	1.6 - 5.3	7.2 - 27.6	*
	EO	0	0 - 3.4	*
G	LEP	*	*	42.5 - 65.9
	FEP	*	*	0.6 - 19.7
	EO	*	*	28.6 - 44.8
H/I	LEP	98.2 - 100	89.6 - 100	*
	FEP	0 - 1.8	0 - 10.4	*
	EO	0	0	*

1 LEP = Limited-English-proficient
2 FEP = Fluent-English-proficient
3 EO = English-Only (native English speaker)

Table 8

Ranges of Ratings of Teachers' English and Spanish Proficiency+, by District and Program

District		Immersion Strategy	Early-Exit	Late-Exit
A	English	4.3 - 4.9	4.6 - 4.9	*
	Spanish	1.7 - 3.4	1.8 - 3.7	*
B	English	5.0	4.8 - 5.0	*
	Spanish	0.6 - 1.1	1.9 - 2.9	*
C	English	4.7 - 5.0	4.3 - 5.0	*
	Spanish	1.0 - 2.2	2.9 - 4.7	*
D	English	*	*	3.4 - 4.1
	Spanish	*	*	2.7 - 4.8
E	English	*	*	3.2 - 4.4
	Spanish	*	*	4.4 - 5.0
F	English	4.0 - 4.4	4.1 - 4.8	*
	Spanish	2.7 - 4.5	2.4 - 3.7	*
G	English	*	*	4.7 - 4.9
	Spanish	*	*	3.1 - 3.8
H/I	English	4.4 - 4.9	4.1 - 4.8	*
	Spanish	3.4 - 4.4	2.4 - 4.8	*

+ Teacher speech samples were rated on a 12-point scale ranging from 0 (for no practical ability to function in the taught language) to 5 (for ability equivalent to that of an educated native speaker).

Schools also vary widely within district and program in the extent to which the teachers have advanced degrees, such as a master's or a doctorate (see Table 9). At one extreme, district E has schools in which two-thirds or more of the teachers in the school have such advanced degrees. In contrast, few teachers have such degrees in the immersion strategy program in districts A, C, and F or in district B's early-exit program.

ERIC

Table 9

Ranges of Percentage of Teachers with Masters or Doctoral Degree, by District and Program

District	Immersion Strategy %	Early-Exit %	Late-Exit %
A	0 - 25.0	0 - 71.0	*
B	0 - 85.7	0 - 33.3	*
C	0 - 33.3	0 - 66.7	*
D	*	*	20.0 - 100
E	*	*	66.7 - 100
F	0 - 20.0	37.5 - 66.7	*
G	*	*	0 - 57.1
H/I	28.6 - 50.0	40.0 - 50.0	*

There is also a great deal of variation within district and program as to the percentage of teachers who have a bilingual or English as a Second Language (ESL) teaching certificate (see Table 10), the percentage of teachers who assign homework (see Table 11), and the percentage of instructional time allocated for each subject (see Table 12).

In sum, schools within districts and districts within programs vary widely on a number of factors that are independent of program that could influence the achievement of students and which must be considered in the data analyses.

Table 10

Ranges of Percentage of Teachers with Bilingual and ESL Certification, by District and Program

District		Immersion Strategy %	Early-Exit %	Late-Exit %
A	Bilingual	33.3 - 100	0 - 100	*
	ESL	11.1 - 100	14.3 - 57.1	*
B	Bilingual	0 - 33.3	0 - 70.0	*
	ESL	0 - 11.1	0 - 12.5	*
C	Bilingual	0 - 100	60.0 - 100	*
	ESL	0 - 100	28.6 - 66.7	*
D	Bilingual	*	*	16.7 - 100
	ESL	*	*	0 - 75.0
E	Bilingual	*	*	75.0 - 100
	ESL	*	*	0 - 46.2
F	Bilingual	100	100	*
	ESL	0 - 100	0 - 66.7	*
G	Bilingual	*	*	83.3 - 100
	ESL	*	*	66.7 - 100
H/I	Bilingual	94.1 - 100	40.0 - 72.7	*
	ESL	33.3 - 100	60.0 - 100	*

Table 11

Ranges of Percentage of Teachers Reporting Homework is Assigned More Than Three Times Per Week, by District and Program

District	Immersion Strategy %	Early-Exit %	Late-Exit %
A	0 - 89.0	0 - 74.0	*
B	0 - 100	0 - 50.0	*
C	11.0 - 100	50.0 - 100	*
D	*	*	33.3 - 100
E	*	*	81.8 - 100
F	60.0 - 100	12.5 - 100	*
G	*	*	28.6 - 100
H	84.0 - 100	50.0 - 100	*

Table 12

Ranges of Percentage of Instructional Time Per Subject, by District and Program

District	Subject	Immersion Strategy %	Early-Exit %	Late-Exit %
A	English Language Arts	47.4 - 65.8	23.6 - 56.0	*
	Spanish Language Arts	0	0 - 43.0	*
	Mathematics	19.0 - 25.0	20.9 - 21.8	*
B	English Language Arts	56.8 - 62.1	28.9 - 64.1	*
	Spanish Language Arts	0	0 - 27.8	*
	Mathematics	16.6 - 22.2	21.0 - 28.2	*
C	English Language Arts	55.3 - 63.7	10.1 - 40.0	*
	Spanish Language Arts	0 - 3.6	31.3 - 42.7	*
	Mathematics	22.7 - 25.4	5.9 - 28.2	*
D	English Language Arts	*	*	22.2 - 49.7
	Spanish Language Arts	*	*	16.9 - 40.5
	Mathematics	*	*	13.5 - 24.5
E	English Language Arts	*	*	19.4 - 36.4
	Spanish Language Arts	*	*	20.2 - 44.0
	Mathematics	*	*	17.7 - 20.2
F	English Language Arts	38.5 - 64.6	36.6 - 53.2	*
	Spanish Language Arts	0 - 10.9	0 - 16.8	*
	Mathematics	14.6 - 30.8	20.1 - 30.8	*
G	English Language Arts	*	*	24.1 - 52.0
	Spanish Language Arts	*	*	5.6 - 43.4
	Mathematics	*	*	17.0 - 24.6
H/I	English Language Arts	59.5 - 69.5	37.8 - 51.6	*
	Spanish Language Arts	0	13.2 - 20.2	*
	Mathematics	21.9 - 25.7	21.4 - 29.5	*

In the nominal program analyses, the programs are viewed as being implemented at the district level. For the immersion program, all five IS/EE districts consistently used nearly all English for instruction. In contrast, there were striking district differences in the pattern of percent English used across the grades for the early-exit and late-exit programs. The differences within district were much less dramatic. (See Chapter I for additional information on the pattern of percent English in the districts.)

In addition to the substantial differences between schools in SES, schools encompass a multitude of attributes that can affect a student's achievement.

<u>**What are the major divisions of these nominal program analyses?**</u>

Because of the necessity to control for district and school when formally comparing programs, the nominal program analyses were divided into three parts, as outlined in Table 4 and Table 5. The first set of analyses compared the immersion strategy and early-exit programs in the four schools that had both programs. In the analyses of these two-program schools, it was possible to estimate the program effect while controlling for the effects of school. Community influences are also controlled in these analyses at the level of the area served by the individual school. The results from these analyses are given in Chapter III.

The second set of analyses compared the immersion strategy and early-exit programs in the schools that had only one of the two programs. In the analyses of these one-program schools, it was possible to control for district. Community influences are controlled at the level of the area served by the entire school district or, in one case, by a pair of nearby school districts. Results from these analyses appear in Chapter IV.

The third set of analyses compared the three districts with late-exit programs. Direct comparison of the late-exit program with the immersion strategy and early-exit programs is not possible, because no district with a late-exit program had either of the other two programs. Comparisons among the three late-exit districts is nonetheless illuminating. These analytic results are presented in Chapter V.

Although statistical comparison of student achievement between the three LE districts as a group and the five IS/EE districts as a group is not appropriate, comparisons among the eight districts (or among the thirteen district-program combinations) can be made less formally through a special descriptive analytic procedure. Any such comparisons must be made in the context of the many differences among the districts. The fourth set of analyses was performed to facilitate these comparisons. Through the use of special graphical techniques, the relative gains of two or more groups can be evaluated. The graphs cannot identify the source of the differences; for example, whether differences stem from student ability or from the educational program. With such a small number of districts that are disparate in so many ways, such attributions cannot be made statistically. These special graphs, called TAMP (Trajectory Analysis of Matched Percentiles) figures, appear in Chapter VI.

<u>What do the Achievement Tests measure?</u>

The Tests of Basic Experiences (TOBE) were given to kindergarten students. With the TOBE, the directions and items are read to the students by the test examiner, and the students indicate their answers by marking one of four pictures. The mathematics tests measures students' mastery of fundamental mathematical concepts and the terms associated with them, as well as understanding of the quantitative relationships between objects. The TOBE language test measures vocabulary, sentence structure, verb tense, sound-symbol relationships, and letter recognition. In addition, there are items measuring listening skills, perception of the meaning of symbols, and derivation of the meaning of "nonsense" words from their context.

The Comprehensive Tests of Basic Skills, Form S (CTBS-S), Level B were administered to students in the fall of grades one through six, and Level C to students in the spring of grades one through six. There are multiple levels of the CTBS-S tests for these grades. At the lowest level most of the test items are read to the students, and as the test level increases the students read more of the items to themselves. While having substantial overlap in content coverage and item difficulty, the adjacent test levels gradually shift their content emphasis to match changes in student development. The mathematics score is based on computation, concepts, and applications. The tests contributing to the first grade language score measure sentence structure and syntax, as well as the use of plurals, verbs, comparative adjectives, and possessive pronouns in Level B, and verb forms, subject-verb agreement, spelling, capitalization, and punctuation in Level C; higher levels measure usage and paragraph organization. The first grade reading tests measure letter sounds, word recognition, and reading comprehension; higher levels measure vocabulary and reading comprehension.

Given that the norms for the TOBE-2 and the CTBS-S were established in 1977-78 and 1972-73, it is appropriate to ask the question of their relevance to students in the current study. Cannell asserted that "standardized, nationally normed achievement tests. . . are inflated and misleading" (Cannell, 1987, p. 3). Motivated by Cannell, Linn, Graue, and Sanders (1990) effected a study of norm-referenced test results reported by states and school districts and of factors related to those scores. They found that the "norms obtained for grades one through eight during the late 1970s or early 1980s are easier on most tests than more recent norms" (p. 13). However, test norms may be changing faster than student achievement. The implications for this study are limited. The performance of students in each of the program models are only being compared to one another. Thus any potential overestimation in performance would likely affect each group equally.

ERIC

What are the minor divisions of the analyses?

The foundation for the analytic comparisons of two-program IS/EE schools (Chapter III), one-program IS/EE schools (Chapter IV), and the three LE districts (Chapter V) is individual growth curve modeling using the HLM computer program (Bryk et al., 1988). However, because CTBS test scores in English are not available for kindergarten, individual growth curve modeling for kindergarten through third grade is not appropriate. The individual growth curve analyses require that all of the test scores be on the same scale. The TOBE language subtest administered to kindergarten students is not on the same scale as the CTBS language arts and reading subtests in English, and it is not appropriate to try to put it on the same scale. Thus the K-1 analyses, comparing the programs at spring of first grade, had to be separated from the 1-3 analyses, comparing the growth curves from first through third grades, at least for the language arts and reading subtests.

The TOBE mathematics subtest administered in kindergarten is not on the same scale as the CTBS mathematics subtest, either. Because of the similarity of content of the TOBE and CTBS mathematics tests, it would be theoretically appropriate to conduct an analysis to produce equated scores for those two tests. Such equating would permit the use of individual growth curve analyses for kindergarten through third grade using the equated scores. While it would be useful to compare such K-3 analyses with the results of the two-part analyses presented here, resource limitations prevented us from performing the K-3 mathematics analyses.

In addition to this division between K-1 and 1-3 analyses, the late-exit analyses also included separate 3-6 analyses. Since no IS/EE 3-6 cohorts were included in the study, there is no comparison group for the LE 3-6 cohorts. Comparisons among the three late-exit districts can be made, however. In addition, combining the 1-3 and 3-6 analyses into 1-6 analyses allows a relatively direct comparison of achievement between these two sets of grades.

Finally, every set of analyses was subjected to a series of sensitivity tests. By varying the model, the subjects included in the analysis, and even the analytic method, it is possible to see the extent to which any conclusions are sensitive to these variations. While literally hundreds of models were estimated, only a small fraction of them are presented, and only the especially important variations are mentioned.

How many students and schools were included in the analyses?

The number of students and schools used varies according to the analysis. Table 13 shows the number of students in each program for the two-program schools analyses. Schools 12 and 14, which are in district B, represent the majority of the students in all of the two-program schools analyses. School 21, in district C, has relatively few students in each program, but all of them had kindergarten test scores (pretest). School 51, in district F, also had relatively few students, none of whom had pretest scores. The absence of pretest scores is a consequence of the fact that the program in this district lost a substantial part of its funding, which resulted in the 113 kindergarten students being exited from the program, and therefore from the study.

Table 13

Number of Students by School and Analysis for Two-Program Schools

		School 12	School 14	School 21	School 51	Total
K-1 Analyses All Students	IS	79	35	17	29	160
	EE	33	30	14	11	88
	Total	112	65	31	40	248
K-1 Analyses Students with Pretest	IS	65	23	17	0	105
	EE	12	22	14	0	48
	Total	77	45	31	0	153
1-3 Analyses All Students	IS	72	28	12	27	139
	EE	27	22	7	11	67
	Total	99	50	19	38	206
1-3 Analyses Students with Pretest	IS	58	16	12	0	86
	EE	6	15	7	0	28
	Total	64	31	19	0	114

Table 14 gives the number of students in each school included in the one-program IS/EE analyses. Because of the very small number of students in certain schools, some of the schools were combined (as indicated in the table) for most analyses. "School 15" is not a physical school, but rather is used to denote a sizable group of students who were in school 13 for kindergarten but in school 14 thereafter.

Table 14

Number of Students by School and Analysis
for One-Program Schools

District	School Group	School	K-1 Analyses: All Students	K-1 Analyses: Students w/Pretest	1-3 Analyses: All Students	1-3 Analyses: Students w/Pretest
A	IS01	01	68	43	59	34
B	IS10	10	12	12	9	9
C	IS2A	22	25	25	18	18
		24	7	1	5	1
H	IS7A	71	4	0	4	0
		75	20	11	19	10
		93	10	0	8	0
H	IS7B	72	15	12	10	8
		73	9	7	8	6
		74	7	3	5	2
		78	13	9	10	6
		79	3	0	3	0
		92	3	0	3	0
H	IS7C	77	9	0	9	0
		91	4	0	4	0
H	IS7D	76	22	12	20	11
A	EE02	02	35	26	31	22
A	EE04	04	13	4	11	2
A	EE05	05	36	25	30	19
B	EE11	11	31	31	22	22
B	EE13	13	25	24	16	15
B	EE15	15	22	22	18	18
C	EE2A	20	40	40	30	30
		23	5	5	4	4
F	EE50	50	24	0	23	0
I	EE81	81	18	13	13	8
I	EE82	82	17	11	14	8
I	EE83	83	35	25	31	21
I	EE84	84	10	0	9	0

Table 15 shows the number of students in each of the late-exit schools. Because district G did not enter the study until spring of the first year, none of the students in district G had pretest scores. This is critical as the analyses require a measure of achievement p. .or to study participation. Of course, none of the students entering the tudy in third grade had pretest scores, either.

Table 15

Number of Students by School and Analysis
for Late-Exit Schools

		K-1 Analyses		1-3 Analyses		
District	School	All Students	Students w/Pretest	All Students	Students w/Pretest	3-6 Analyses
D	30	42	21	35	16	21
E	40	13	13	4	4	5
	41	29	28	26	25	25
	42	0	0	0	0	8
	43	28	26	17	16	14
	44	23	21	20	18	19
	45	20	19	16	16	10
	46	4	4	3	3	17
G	60	10	0	9	0	4
	61	5	0	4	0	2
	62	14	0	11	0	14
	63	20	0	15	0	5
	64	3	0	3	0	9
	65	9	0	7	0	1

How were the kindergarten test scores used in the analyses?

Kindergarten students were administered Level K of the TOBE 2 test, referred to hereafter simply as the TOBE. Although the spring kindergarten TOBE was administered in twice, once in Spanish and once in English, the fall kindergarten TOBE was administered once, in the child's dominant language. For the vast majority of children, this meant the fall kindergarten TOBE was administered in Spanish. It would be questionable

ERIC

to include a test score in Spanish as part of a growth curve where the other scores were in English. However, the fall kindergarten TOBE, administered in Spanish, may be a reasonable measure of a student's "school readiness." That is, it may be a useful predictor of the student's achievement, even as measured in English. Accordingly, the fall kindergarten Spanish TOBE subtest scores were used as predictors of achievement at spring of first grade, as measured by the English CTBS achievement tests.

What form of the CTBS achievement subtest scores were analyzed?

The achievement tests administered beginning in first grade were from the appropriate level (B, C, 1, 2, or 3) of the Comprehensive Tests of Basic Skills, Expanded Edition, Form S (CTBS). The Expanded Scale Scores (ESS) from the CTBS tests, administered in English, were analyzed. The ESS form a continuous scale from kindergarten through high school. For any one of the three subtests (mathematics, language, or reading), the ESS from any administration of the subtest may be compared with the ESS from any other administration of that same subtest. The ESS are on an equal-interval scale, so a difference of one point has the same meaning regardless of the magnitude of the scores. That is, a change from an ESS of 200 to an ESS of 210 on a subtest has the same meaning as a change from an ESS of 350 to an ESS of 360. The ESS should not be compared across subtests, however. An ESS of 300 for the language subtest does not mean the same level of achievement as a 300 for the mathematics or reading subtest.

Why were CTBS Expanded Scale Scores used?

Using a single continuous scale like the Expanded Scale Score of the CTBS is essential for modeling growth. Therefore the 1-3 analyses needed to be performed on the ESS. For convenience of comparison of the 1-3 results and the K-1 results, the ESS were also used for the K-1 analyses. Thus the outcome measures for the K-1 analyses, the spring first grade ESS for each of the three CTBS subtests, correspond to the initial time point for the 1-3 analyses.

How do national samples perform on the CTBS achievement tests?

National norms are available for the CTBS achievement tests. As these norms are based on national samples of students, they reflect the achievement of students whose primary language is usually English. While the applicability of these norms to describe the academic growth of language-minority students is questionable, these norms are the criteria used to determine whether a language-minority student is limited-English-proficient and therefore in need of special supportive services. Additionally, the norm information permits some understanding of how the distribution of ESS changes from grade to grade.

Table 16 shows the mean and standard deviation of the ESS for the norming sample by grade for each subtest, taken from the CTB/McGraw-Hill CTBS/S Technical Bulletin No. 1 (1974). Two trends can be seen in this table. First, the difference between the means for successive grades diminishes as one moves from the lower grades to the higher grades. Second, the standard deviation increases with increasing grade.

Table 16

Norming Sample:
ESS Means and Standard Deviations by Grade and Subtest

	Mathematics		Language		Reading	
Grade	Mean	S.D.	Mean	S.D.	Mean	S.D.
Spring 1st	259	36.5	283	55.7	263	45.2
Spring 2nd	308	43.0	359	66.0	328	62.2
Spring 3rd	364	54.7	410	75.7	380	70.8
Spring 4th	394	60.1	440	77.3	422	78.8
Spring 5th	436	69.4	470	79.1	459	84.9
Spring 6th	464	80.8	487	84.6	485	93.6

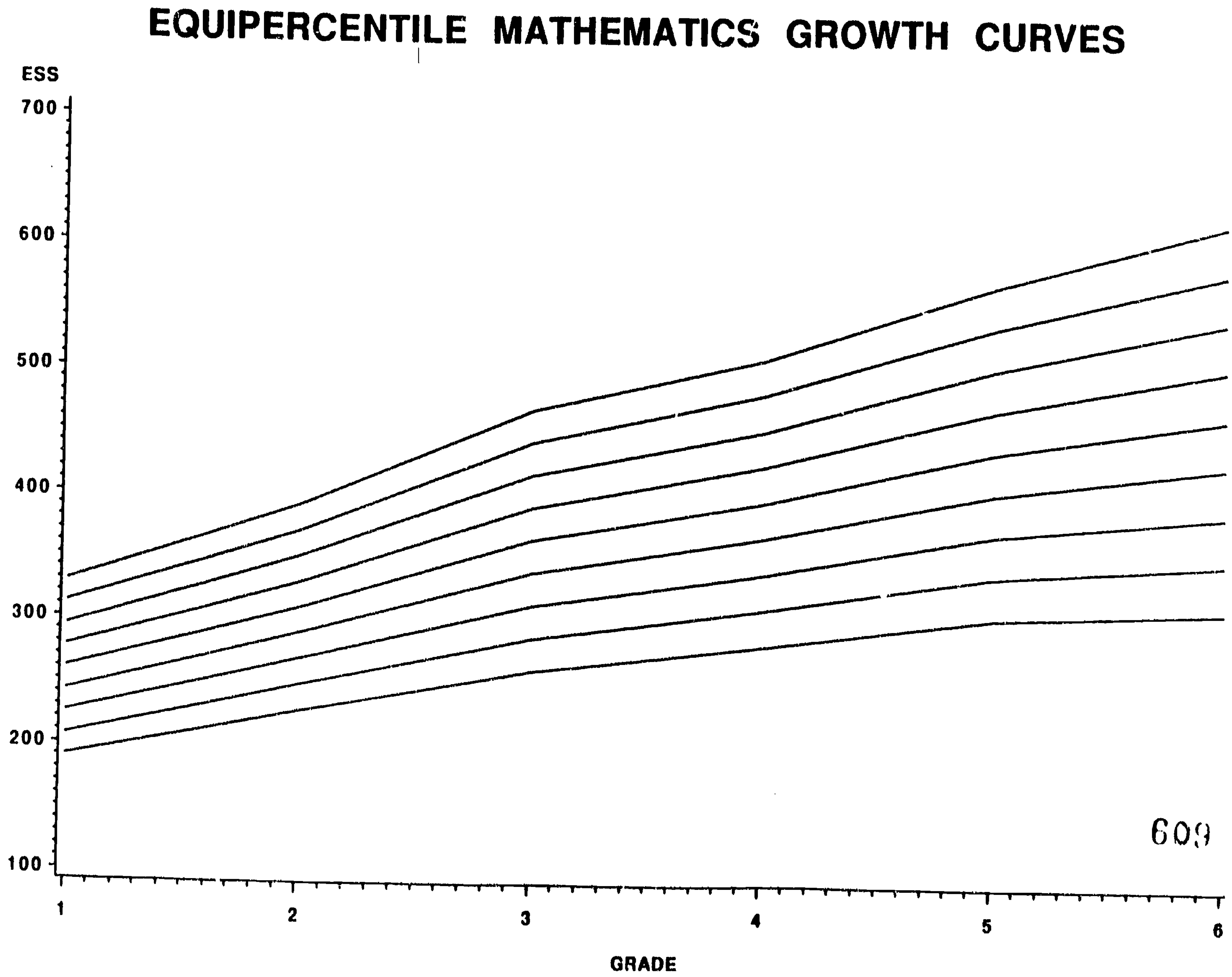
Figure 13 through Figure 15 show "equipercentile growth curves" for spring of first grade through spring of sixth grade for each subtest. The nine curves in each figure connect the 10th, 20th, ..., 90th percentile points from the norming sample at each grade. Thus each curve represents the scores that a student would have to attain to remain at the same percentile. It should be emphasized that these curves do not represent the growth trajectory of any individual. They are based on the cross-sectional norming sample and therefore suffer from all of the flaws inherent in judging student growth using cross-sectional samples (see Chapter VI for further discussion of this point).

Because the ESS distribution is approximately symmetric at each grade, the median score and the mean score are nearly equal. Thus the center curve in each figure, which tracks the 50th percentile (median) score, is nearly the same as the curve that would be obtained by graphing the mean scores given in Table 16.

The same trends noted for the means and standard deviations in Table 16 can be seen in the three figures. The curves tend to flatten at the higher grades; this corresponds to the diminishing differences in mean scores. Also, the distance between the curves tends to increase at the higher grades; this corresponds to the increasing standard deviation.

ERIC

Figure 13



66

608

609

Figure 14

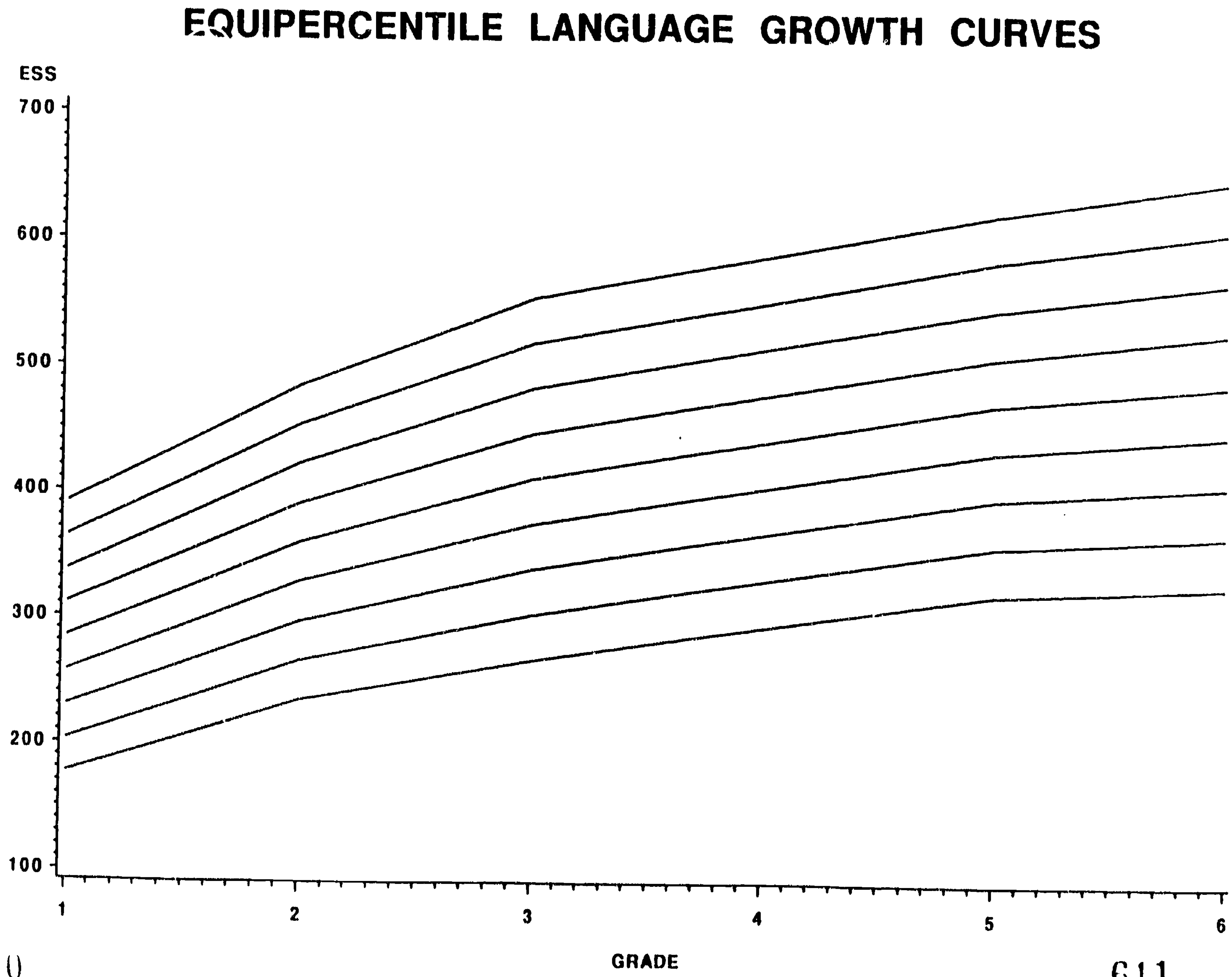
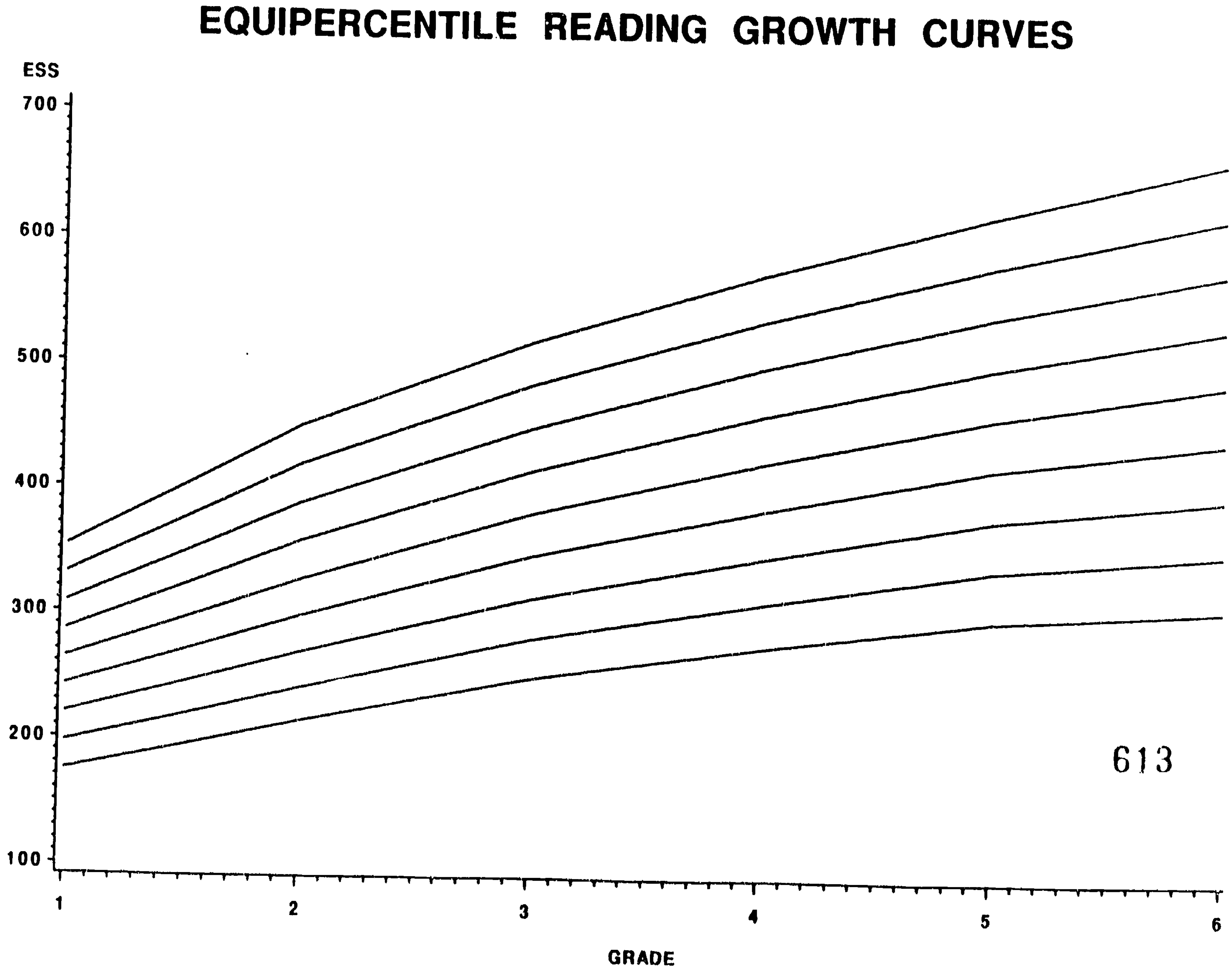

EQUIPERCENTILE LANGUAGE GROWTH CURVES
ESS
700
600
500
400
300
200
100
1
2
3
4
5
6
GRADE

Figure 15



612

613

What measure of time was used in the evaluation of growth?

To evaluate growth, a measure of time is needed. The date of each administration of the achievement tests was known for each student. These dates formed the basis for the time metric. Time was measured in years of schooling, with the "zero" point taken to be spring of first grade. The advantage of setting the "zero" to spring of first grade is that the estimated "intercept" for the 1-3 growth curve analyses is the same as the end point of the K-1 analyses.

The details of the calculation of time are as follows. Usually students were tested some time from March through June, so the midpoint of that time period, May 1, was used. The number of days between the test date and May 1 was divided by 365 to give the fraction of a year. To this was added the number of years the child had been in the program prior to the current school year, reduced by 1. To be included in the study a child had to begin the program in fall of kindergarten, so the number of years in the program is the same as the number of years since entering kindergarten.

A student advancing normally through the grades would have a value of 0.0 on May 1 of first grade, 1.0 on May 1 of second grade, 2.0 on May 1 of third grade, 3.0 on May 1 of fourth grade, and so on. A student who had repeated kindergarten, however, would have a value of 0.0 on May 1 of the second year of kindergarten, a value of 1.0 on May 1 of first grade, 2.0 on May 1 of second grade, and so on.

How were program differences in spring of first grade evaluated?

In the K-1 analyses, program differences in spring of first grade were evaluated using Analysis of Covariance (ANCOVA), controlling for school and various predictors of achievement (covariates). It is desirable in this context to use as one of the covariates a direct measure of each student's achievement or aptitude before entering the program. Since all study students entered the program in kindergarten, test results from fall of kindergarten are usable as "pre-treatment tests" or "pretests." Any test scores after the fall kindergarten scores were possibly affected by the program, so they are not suitable for use as a pretest.

Unfortunately, no fall kindergarten tests were available for many students. This is primarily a consequence of the study design. Some students, and some entire districts, did not enter the study until spring of the first year, so the kindergarten students were not tested until spring. In addition, some students entered the study in first grade (even though all students in the study entered the program in kindergarten). In addition, there is a small amount of missing data among those students who should have been tested in fall kindergarten.

Because so many students lack pretest (fall kindergarten) scores, the K-1 analyses were performed in three parts. First, various covariates other than pretest were used as predictors of achievement at spring of first grade, as measured by the CTBS achievement tests. Second, those covariates were evaluated as predictors for the students who had a pretest to see how different those students were from the larger group. Third, the effect of including the pretest as a covariate was evaluated on this subset of students who had pretest scores. (See Appendix B for a list of covariates included in the analyses by label and name.)

Were test date differences accommodated in the K-1 analyses?

Although all of the students in the K-1 analyses for two-program schools were tested in spring of first grade, it would be desirable for the differences in test date to be accommodated in the analyses. Otherwise, students tested relatively late in the school year would be likely to show higher levels of achievement than those tested earlier in the school year. Since students within a school tended to be tested at about the same time of year, differences in test date could produce apparent school differences where none existed. A systematic difference in test dates for students in the two programs in a school could produce

ERIC

apparent program differences where none existed, or obscure actual program differences.

Unfortunately, the test date was almost completely confounded with the school and study year. That is, each year almost all of the first grade students within a school were tested at about the same time. This creates no problem for the 1-3 analyses, where most of the effect of time is estimated from the year-to-year changes in test scores. The K-1 analyses, however, used only one test score for each student. To the extent that the first grade test dates varied systematically by school or cohort, any estimated effects of the test dates would be likely to reflect school or cohort differences rather than actual test date differences.

Nonetheless, an attempt was made to include test date as a predictor in the K-1 analyses. The results were extremely unstable and often indicated a substantial negative effect of time. That is, students who took the test later in the year often did less well on the achievement test. An evaluation of this phenomenon showed that the effect within school was usually positive, so that between-school differences in test date accounted for the negative effect. That is, when students from certain schools were tested later in the year, they tended to score lower than when tested earlier in the year. It is unclear why this was so. Where the within-school effect was negative, it was found to reflect cohort differences. Even a small cohort difference could produce sizable time effects. For example, if the first grade students were tested on May 1 in the first year of the study but, in the second year of the study, the first grade students were tested on May 14, and the first grade students in the first year averaged just 1 point higher on a test than the first grade students in the second year, it would be estimated that students lose 1 point every 13 days, or about 20 points a year. As a consequence, the date of the spring first grade test (as measured on the time metric described above) was not included as a predictor of achievement in the K-1 ANCOVA analyses.

How were variables selected as potential predictors of achievement?

For all of the analyses in this report, the set of possible predictors of achievement was restricted to a relatively small number by considering only variables unlikely to be affected by the program. At first glance, including variables likely to be affected by the nominal program could have obscured the program effect. However, there are situations in which including some variables that are possibly affected by the program is reasonable (Rosenbaum, 1984).

Avoiding variables likely to be affected by the program meant eliminating any variables associated with the child's educational experience, including district, school, or teacher attributes or behaviors. It also meant eliminating any student or parent behaviors that were likely to have been influenced by the child's educational experience, such as whether the child does homework or the amount of time the child spends engaged in various non-academic activities such as reading or watching television.

An exception to this exclusion rule is that the number of days absent was included even though absences might have been influenced by the child's educational experience. The average number of absences is both an indicator of whether the child actually received the treatment and an indicator of the child's health. It was therefore considered too important to be omitted as a predictor.

Although the number of days absent was the only variable included as a possible predictor that was likely to be affected by the program, two variables that are possibly affected by program were chosen as potential predictors: the number of books in the home, and whether the parents use English when speaking to each other. The number of books in the home (excluding school books) might have been affected by programs that strongly encouraged or even subsidized book purchases. However, there was no evidence of such activities by the schools and districts in this study. The amount of Spanish and English used at home between parents might have been indirectly affected by the program through the student's use of English. A student receiving little instruction in English might be less likely to use English at home, and this might make the parents less likely to use English with each other. While there is a strong correlation between the amount of English used between parents and the amount of English used between children and parents, the correlation was not as high as expected. Table 17 gives the correlation matrix for three variables: ANYEPTOP (Was any English used Parent-To-Parent?), ANYEPTOC (Was any English used Parent-To-Child?), and ANYECTOP (Was any English used Child-To-Parent?). The correlations among the variables range from about .44 to about .61. The three-way frequency distribution of these variables provides additional information on the relationship among these variables (see Table 18).

Table 17

Correlations for Amount of English Used at Home
(N = 1164)

	ANYEPTOP	ANYEPTOC	ANYECTOP
ANYEPTOP	*	0.6071	0.4437
ANYEPTOC	0.6071	*	0.5257
ANYECTOP	0.4437	0.5257	*

Table 18

Three-Way Frequency Distribution for Amount of English Used at Home

ANYEPTOP	ANYEPTOC	ANYECTOP	Frequency	Percent
No	No	No	583	50.1
No	No	Some	198	17.0
No	Some	No	52	4.5
No	Some	Some	118	10.1
Some	No	No	9	0.8
Some	No	Some	11	0.9
Some	Some	No	12	1.0
Some	Some	Some	181	15.5

After restricting the possible predictors of achievement in this way, the number of predictors was further reduced statistically. Variables were tested for statistically significant associations with initial status or growth rate. Where several variables were highly associated, a few from the group were selected for further analysis, and a tentative model was developed using these few variables. Next, the omitted variables were tested again for statistical significance. Through this iterative process, models were developed for predicting the initial status and growth rate for each of the achievement subtests. Thus, some covariates will show up in the model for one subtest but not in the model for another subtest. Some variables that were not statistically significant were included in these models to facilitate comparisons with other models in which the variables were statistically significant. For convenience of reference, these variables as a group will be called "background variables," even though this term is not universally applicable.

What were the characteristics of the predictors of achievement?

The characteristics of the background variables used as predictors of achievement in the analyses are summarized in Table 19, Table 20, and Table 21. These tables are based on all 1,164 students included in the analyses. Relatively few of these students had missing data for any of these variables because all students included in the analyses were required to have parent interview data available. The small amounts of missing data were imputed by using relationships with other parent interview data.

Table 19

Descriptive Statistics for Covariates
(N = 1164)

Variable	Mean	S.D.	Min	Max
EDAVG	7.734	3.467	0	19
PRESCHY	0.369	0.483	0	1
FEMALE	0.516	0.500	0	1
OCAVG	3.716	1.627	1	10
AGEMOTH	2.051	0.417	1	3
BOOKSHM	1.942	1.418	0	4
ANYEPTOP	0.183	0.387	0	1

Table 19 provides the mean, standard deviation, minimum, and maximum for the seven background variables used in the analyses, while Table 20 gives frequency distributions for the five covariates with restricted ranges. EDAVG is the parents' average years of education, derived from questions 61 and 62 on the parent interview. PRESCHY has the value 1 if the student attended preschool and 0 otherwise (question 17 on the parent interview). FEMALE has the value 1 if the student is a girl and 0 if the student is a boy, obtained from the student database. OCAVG is the average of the socioeconomic occupation index for the parents, derived from questions 63 and 64 on the parent interview. AGEMOTH is a three-valued variable indicating the mother's approximate age (1 = under 25, 2 = 25 to 34, 3 = 35 and over), derived from parent interview question 7. BOOKSHM is the number of books in in the home, not counting schoolbooks, in five categories (0 = None, 1 = 1 to 10, 2 = 11 to 20, 3 = 21 to 30, 4 = 31 or more). It is derived from question 37a on the parent interview. ANYEPTOP has the value 1 if any English is used Parent-To-Parent and 0 if only Spanish is used, as described above. It is derived from question 34 on the parent interview.

ERIC

Table 20

Frequency Distributions for Selected Covariates

PRESCHY	Frequency	Percent
0: No	735	63.1
1: Yes	429	36.9

FEMALE	Frequency	Percent
0: Male	563	48.4
1: Female	601	51.6

AGEMOTH	Frequency	Percent
1: Under 25	73	6.3
2: 25 - 34	959	82.4
3: 35 & Over	132	11.3

BOOKSHM	Frequency	Percent
0: None	185	15.9
1: 1 - 10	376	32.3
2: 11 - 20	199	17.1
3: 21 - 30	130	11.2
4: 31 & Over	274	23.5

ANYEPTOP	Frequency	Percent
0: No English	951	81.7
1: Some English	213	18.3

The correlation matrix for all seven covariates is given in Table 21. The correlations among the three variables that indicate socioeconomic status (EDAVG, OCAVG, and BOOKSHM) range from about .23 to about .37. The variable ANYEPTOP has correlations with the three SES variables ranging from about .10 to about .20, indicating that higher SES is somewhat associated with greater likelihood that the parents use some English when speaking to each other. The correlation between AGEMOTH and EDAVG is -.14, showing an association between more years of parents' education and younger mothers. No other correlation is as large as .07 in magnitude.

Table 21

Correlation Coefficients for Covariates
(N = 1164)

	EDAVG	PRESCHY	FEMALE	OCAVG	AGEMOTH	BOOKSHM	ANYEPTOP
EDAVG	*	0.0440	0.0227	0.3373	-0.1391	0.3692	0.1969
PRESCHY	0.0440	*	0.0053	-0.0334	0.0353	0.0466	0.0437
FEMALE	0.0227	0.0053	*	0.0024	0.0063	0.0559	0.0134
OCAVG	0.3373	-0.0334	0.0024	*	0.0241	0.2294	0.0952
AGEMOTH	-0.1391	0.0353	0.0063	0.0241	*	0.0559	-0.0682
BOOKSHM	0.3692	0.0466	0.0559	0.2294	0.0559	*	0.1857
ANYEPTOP	0.1969	0.0437	0.0134	0.0952	-0.0682	0.1857	*

How were the individual growth curves estimated?

For the 1-3 analyses, first grade through third grade information was used to estimate individual growth curves. For each student with test scores at two or more time points, an individual growth curve was fit to the subtest scores for each subtest. The individual growth curves fitted were curved lines. Although straight lines could have been used, the national norms indicate that the amount of growth can be expected to decline from year to year. That is, a student scoring at the same percentile every year can be expected to show greater growth in the early grades than in the later grades. Ignoring this curvature could produce misleading results in the comparison of students with different patterns of data. Students with only first and second grade test scores, for example, would be expected to show greater growth rates than students with first, second, and third grade test scores because of the flattening of the growth trajectories implied by the national norms.

With only three time points for each student, however, any individual growth curve other than a straight line would be guaranteed to fit the data exactly. This would result in no estimate of variability, and the resulting growth curves could not be used to make inferences about differences among students. Therefore, the amount of curvature in the growth could not be allowed to vary at the level of the individual child. Instead, curvature had to be assumed to be the same for all children, or at least the same for groups of children.

However, a slope and an intercept could be estimated separately for each individual, using the achievement subtest scores and the date the test was administered. The slope is interpreted as the average number of ESS points the child increased each year of school, or the child's growth rate. The intercept is interpreted as the initial status of the child -- the ESS of the child at spring of first grade.

In the terminology of hierarchical linear models (HLM), the slope and intercept parameters are called random effects because they are permitted to vary from individual to individual. The coefficient of the curvature is called a fixed effect because it has the same value for all individuals within a group. The slope, intercept, and coefficient for curvature together comprise the within-child model.

<u>**In what sense is the model hierarchical?**</u>

The model developed using the HLM program is called hierarchical because the random effects (the slope and intercept) from the within-child model can be predicted in a between-child linear model. That is, information that can be associated with a child but which does not vary over time can be used to predict the growth rate (slope) and initial status (intercept) for a child. This produces a two-level hierarchy of linear models: one model for predicting an individual's test scores at each time, and a separate model for predicting the random effect parameters of the first model. These two linear models are called the within-child model and the between-child model, respectively. By including a variable indicating program among the predictors of the initial status and growth rate, it is possible to estimate program differences in initial status and growth rate. By including additional background information associated with the child, such as parents' education, the estimated program effects are adjusted for that additional information.

Recently, the hierarchical linear model has been extended to three levels, permitting the simultaneous modeling of individual growth and school effects (Bryk and Raudenbush, 1988). In the context of this study, the three linear models are (1) a model for predicting an individual's test scores at each time (the within-child model), (2) a model for predicting the initial status and growth rate for each child in each school (the between-child, within-school model), and (3) a model for predicting the average initial status and average growth rate in each school (the between-school model). This differs from the two-level hierarchical model in that schools are explicitly included in the model, with the average initial status of the students in a school and the average growth rate of those students treated as random effects. The three-level version of the HLM computer program is still under development, and it was not feasible to use it for all of the analyses where the three-level model would be appropriate. However, it was possible to estimate some three-level models in connection with the 1-3 analyses for the one-program schools, where the ability to treat both students and schools as having random effects is most important (see Chapter IV).

<u>How is the coefficient for curvature interpreted?</u>

Unlike the intercept (initial status, at spring of first grade) and the slope (the growth rate from spring of first grade to spring of second grade), the curvature has no natural interpretation. There are several equivalent codings for the curvature variable. They are equivalent in the sense that changes in the estimated coefficients for intercept, slope, and curvature permit any of the codings to produce the same growth curves as any other coding. The coding used was chosen for both analytic and interpretation reasons.

ERIC

The curvature variable was constructed as a quadratic in time that was (approximately) orthogonal to the intercept and slope. For the 1-3 analyses, the values of the time variable T are approximately 0, 1, and 2 while the intercept variable is always 1. The regression of T-squared (TSQ) on T and an intercept for T equal to 0, 1, and 2 produces the equation TSQ = 2 x T - 1/3. Thus the unique contribution of TSQ can be given by the variable Q = TSQ - 2 x T - 1/3 (where Q stands for quadratic). For T equal to 0, 1, and 2, Q has the values 1/3, -2/3, and 1/3. Since these values for the variable Q have a mean of zero, the coefficient of Q does not affect the coefficient of the intercept. And since these values for Q have a slope of zero, the coefficient of Q also does not affect the coefficient of the slope.

Because the times used in the analyses were not exactly 0, 1, and 2, and because not all students in the 1-3 analyses had all three time points, the curvature variable Q is somewhat correlated with the intercept and the slope. However, the correlations are small and make the model-fitting using Q much more stable than if TSQ were used. In addition, it permits the interpretation of the coefficient of Q as representing essentially pure curvature. A positive coefficient for curvature means that growth has an upward curve (the growth gets steeper), whereas a negative coefficient means that growth has a downward curve (the growth gets flatter).

In analyses where the coefficient for curvature is the same for all students, the program comparison may be made on the basis of the coefficients of program as a predictor of intercept (initial status) and slope (growth rate). Even if the coefficient for curvature is found to differ according to one or more background variables, the essential information about program differences is encompassed in the initial status and growth rate prediction equations. However, in models where the curvature is found to differ significantly according to the program, that curvature difference must be considered when comparing the programs.

How were student absences accommodated in the model?

All models include an adjustment for the effect of student absences from school. For the ANCOVA models predicting first grade test scores, the number of absences in first grade was included as a covariate. Where the number of first grade absences was not available, the average number of absences in second and third grade was substituted. This variable was included in the models that had any covariates regardless of its statistical significance.

Growth curve models included the average number of absences for the grades covered by the model. This average absence variable was always included as a predictor of initial status (intercept) regardless of statistical significance. This allows the entire growth curve to be shifted vertically to accommodate the average number of absences. In addition, the average absence variable was included as a predictor of growth rate (slope) in any model where it was statistically significant. This allows the growth curve to be tilted, which results in a model in which absences have a cumulative effect.

An alternative to using the average number of absences as a predictor of initial status is to use the actual number of absences during a year to predict the test score for that year. If every child were absent exactly the same number of days every year, the two formulations would be equivalent. However, the year-to-year variations in the number of absences might help improve the prediction of year-to-year variations in test scores. Unfortunately, the number of absences was often missing for individual years even though it was essentially never missing for all of the years included in the growth model. The HLM program allows the use of within-student covariance matrices calculated by the pairwise-missing method, permitting the inclusion of students with missing absence data. Unless the absence data is missing completely at random, however, the results can be misleading. Nonetheless, some models were estimated using the number of absences as a within-student variable (using the pairwise-missing capability) instead of using the average number of absences as a predictor of initial status. These models were not materially different from each other.

How were missing test scores accommodated?

Test scores are missing for one of three reasons: design, attrition, or because the student was still in the study but no test score was obtained. By design, test scores are missing for students before they entered the study or after the study ended. The design also includes fall testing in English only in the year the student entered the study. In fall kindergarten, testing was not done in English unless it was the student's dominant language. (Relatively little testing was done in Spanish after kindergarten, but Spanish achievement is not addressed in this report.)

Attrition of the student from the study meant no additional test scores could be obtained for that student. The predominant reason for attrition was that the student's family moved outside the district. Students who were mainstreamed were not considered to have left the study: test scores for mainstreamed study students are not missing by design.

Occasionally test scores are missing for students who were still in the study that year. The typical reason that a student is missing one or more subtest scores in a given year is that the student was absent on the day the subtest was administered. For the analyses, a student's test scores for a year are included only if all three subtests were available (both subtests in the case of the TOBE tests administered in kindergarten). This ensures that exactly the same students are included in the analyses of the mathematics, language, and reading subtests.

After reducing the test scores to complete sets, the patterns of available data were evaluated. For any given analysis, only a handful of students were missing test data between two available test scores. All of the other students had an uninterrupted series of test scores. This includes students who lacked earlier or later scores by design, attrition, or by happenstance.

The few students who had second and third grade test data but lacked first grade test data were excluded from the 1-3 analyses. Because any student with first grade test data was included in the main K-1 analyses, this ensures that the students in the 1-3 analyses are a subset of the students in the main K-1 analyses. Most of the students in the main K-1 analyses were also includible in the 1-3 analyses because they had either second or third grade test scores or both. No student had only first grade test scores by design. Some students had only kindergarten, first grade, and second grade test scores by design, however.

Propensity modeling, described below, accommodates students omitted entirely from analyses due to missing data. Students included in the analyses but with less than complete test score data were a potential source of bias. Students who left the study early might have relatively high or relatively low achievement, for example. Such possible attrition biases were checked for by comparing models with and without attriters and by relating residuals to data patterns. Unlike the situation in later grades, where students who leave school are likely to be those who are not performing well, the attrition among these young students appears not to be systematically related to achievement.

How were student differences between programs accommodated in the analyses?

The inclusion of "propensity scores" in the IS/EE analyses is designed to adjust the estimated program effect to account for differences between the children in the immersion strategy program and those in the early-exit program. Such differences generally arise either because of a selection effect, or because of an attrition effect, or both.

As the term "selection" is used here, a selection effect can exist even if parents do not explicitly select a bilingual program. In fact, it

ERIC

is unlikely that parents make any explicit choice of bilingual program. Whenever the children in the two programs differ when they begin the program (i.e., initial differences), there is a selection effect. Because most of the schools in IS/EE districts offer only one program, differences in the neighborhoods served by schools are likely to appear as selection effects. In addition, there is the opportunity for school district administrators to choose which schools will implement which program, thereby explicitly selecting programs for the associated neighborhoods. There is evidence that some districts deliberately implemented immersion strategy programs in their "best" or their "worst" schools.

Even if there is no selection effect, so that the students in the two programs are comparable when they enter the programs in kindergarten, there may be an attrition effect. While the main cause of attrition in the student population studied is that the family moved out of the district, such attrition could create differences between the programs. For example, if a plant closing causes many laborers to move away from one neighborhood but has relatively little effect on the laborers in another neighborhood, a socioeconomic difference between the neighborhoods might be created where none existed before.

Regardless of the origin of differences between the students in the two programs, any systematic differences can have a profound effect on conclusions about program effects. For example, if the children in the immersion strategy program were advantaged relative to the children in the early-exit program, and if immersion strategy students were found to have higher achievement than early-exit students, concluding that IS was a superior program might be incorrect. After adjusting for the relative advantage of the immersion strategy students, the program effect might be reversed, resulting in the opposite conclusion: that the EE program was superior.

It is possible to assess the extent to which the program effects might be due to systematic differences in the children in the two programs. One way is to include various covariates as predictors of

ERIC

initial status and of growth rate and then to consider the resulting changes in the program effect. A more direct method is to determine the ways that students in the two programs differ from each other and use that information to develop a "propensity model" that estimates the propensity, or tendency, of a child with certain attributes to be found in one program or another. It can be shown that including the estimated propensity score in the model adjusts the estimated program effect to account for all of the differences incorporated in the propensity model (Rosenbaum and Rubin, 1983).

Because the propensity model needs to reflect both selection effects and attrition effects, the propensity score needs to be estimated separately for each analysis. That is, once we had determined which students could be included in an analysis, we then examined that set of students for any demographic features that make them or their program distinctive. This was done for each of many possible covariates.

What variables were considered for the propensity models?

All variables that were inherently program-related were excluded from consideration for the propensity models. Also excluded were variables that varied at the district level, since district effects were to be modeled individually. This left some school variables and many parent-related variables.

Preliminary analyses to reduce the number of candidate variables were successful in reducing the number of school variables to eight and the number of parent-related variables to thirty. The eight school variables all related to the socioeconomic status (SES) of the families served by the school. Three income variables were included: the percentage of students who were from homes of low income, the percentage from low-middle income homes, and the percentage from middle income or above homes. Four variables relating to the occupation of the head of household were included as candidates: the percentage unemployed, the percentage in unskilled labor, the percentage in blue collar occupations, and the combined percentage clerical, white collar, or professional. The final variable was the percentage of the students in the school who received free or reduced-price meals.

Although the propensity modeling operates at the level of the individual student, school variables must be tested at the school level. (A school-level variable evaluated at the student level would have artificially low within-program variability because all of the students in a school would have been given exactly the same value for the variable.) There were 24 one-program schools in the IS/EE districts for which the school site fact sheet data were available. Each of the eight variables was checked for systematic differences between the 11 immersion strategy schools and the 13 early-exit schools. In addition, an overall Multivariate Analysis of Variance (MANOVA) test was performed on the eight school variables as a separate subset.

The programs were compared on school variables without adjusting for district effects (equivalent to ordinary two-sample t-tests) and again after adjusting for district mean effects using Analysis of Variance (ANOVA). In addition, district by program interactions were tested. None of these statistical tests achieved significance even at the .10 level. In fact, most of the F-ratios were less than 1, indicating that the school-to-school variability within the programs was greater than the variability between the programs (although not significantly greater).

Because the number of students in the study varies considerably from school to school, analyses weighted by the number of students were undertaken to ensure that this imbalance did not introduce program differences on any of the eight school SES variables. Even if the schools in the two programs have essentially the same levels of SES, as indicated by the unweighted analysis described above, if in one program the high-SES schools have relatively more students, but in the other program the low-SES schools have relatively more students, the two programs will not be comparable on school SES. Accordingly, analyses weighted by the number of study students in the school were performed.

Table 22 summarizes the results of those weighted analyses evaluating program effects. Only the "% low-middle income" variable reaches statistical significance at the .05 level and only after adjusting for district effects. Since this is the only test of the dozens performed that is statistically significant, and neither the overall MANOVA test nor the corresponding unweighted test are significant, it is reasonable to ascribe it to chance. That is, it is reasonable to conclude that the two programs do not differ in school SES, even after taking into account the number of students included in the study in each school. To the extent there was any difference, the immersion strategy schools had a slightly higher proportion of low income and slightly lower proportions of low-middle and middle and higher incomes.

Table 22

Program Differences on School SES Variables
Weighted by Number of Study Students

	Unadjusted		Adjusted for Program	
Variable	F-Stat	p-Value	F-Stat	p-Value
% low income	2.54	.1256	3.79	.0675
% low-middle	3.05	.0948	5.28	.0337
% middle & up	0.83	.3732	1.27	.2744
% unemployed	1.61	.2172	0.51	.4830
% unskilled labor	0.07	.7916	0.53	.4776
% blue collar	0.20	.6571	0.00	.9548
% clerical, white collar, professional	1.05	.3175	1.58	.2249
% free or reduced price meals	1.92	.1793	3.16	.0926
Overall MANOVA (Wilks' Criterion)	1.49	.2403	1.16	.4006
(exact d.f.)	(8, 15)		(8, 11)	

District differences on the school SES variables were also analyzed, both unadjusted and after adjusting for program differences. Table 23 summarizes the results of the analyses weighted by the number of students in each school. Only the percent of unskilled laborers shows a district effect significant at the .05 level. It was somewhat surprising not to find many district differences. The near-absence of district differences leads to the conclusion that school SES differences among the one-program schools within a district may be as great (or as small) as the district-to-district differences. This indirectly indicates that school effects should not be neglected in the analyses.

Table 23

District Differences on School SES Variables
Weighted by Number of Study Students

Variable	Unadjusted F-Stat	Unadjusted p-Value	Adjusted for Program F-Stat	Adjusted for Program p-Value
% low income	0.34	.8465	0.74	.5800
% low-middle	1.21	.3378	1.92	.1510
% middle & up	0.40	.8030	0.54	.7092
% unemployed	2.27	.0993	1.88	.1587
% unskilled labor	4.14	.0141	4.14	.0150
% blue collar	0.95	.4549	0.86	.5089
% clerical, white collar, professional	0.51	.7309	0.67	.6201
% free or reduced price meals	0.78	.5525	1.16	.3606
Overall MANOVA (Wilks' Criterion)	1.25	.2392	1.17	.3146
(approx. d.f.)	(32, 45.85)		(32, 42.16)	

The apparent district differences in the percent of unskilled laborers should not be relied upon as an indication of district SES. Indeed, the absence of statistically significant differences among the districts on the other income and occupation variables as well as the non-significance of the overall MANOVA test indicates that the districts do not differ appreciably in school SES. To the extent that there are district differences in the proportion of unskilled laborers, districts A, C, and H/I all averaged about one-third, but districts B and F averaged about two-thirds. Again, the absence of statistically significant differences among the other occupation percentages (which necessarily add to 100 percent for each school and, therefore, for each district) implies that the five districts do not differ in the overall SES of their schools.

What were the models of propensity?

The propensity models were based on the parent interview. After reducing the eligible variables to 30 variables based on preliminary results, logistic regression was used to establish a model for predicting membership in the immersion strategy program. Table 24 presents the estimated coefficients for the logistic regression propensity models for the K-1 analyses and the 1-3 analyses of one-program schools. A positive coefficient indicates that higher values of the variable are associated with a greater propensity to be in the immersion strategy program. A negative coefficient indicates that higher values of the variable are associated with a greater propensity to be in the early-exit program.

Table 24

Propensity Score Logistic Regression Models
for One-Program K-1 and 1-3 Analyses

Variable	PONEK1	PONEK1PN	PONEK1PP	PONE13	PONE13PN	PONE13PP
INTERCEPT	5.679	4.907	6.138	5.399	4.110	6.619
DISTA	-0.745	-0.575	-0.493	-0.738	-0.576	-0.468
DISTB	-2.565	-2.281	-2.367	-2.504	-2.189	-2.356
DISTC	-0.802	-0.741	-0.770	-0.771	-0.603	-0.764
EDAVG	-0.123	-0.111	-0.106	-0.126	-0.114	-0.107
PRESCHY	0.403	0.375	0.386	0.669	0.838	0.933
BOOKSRD	0.046	0.049	0.051	0.048	0.052	0.055
RPTOCLSY	-0.643	-0.736	-0.719	-0.698	-0.900	-0.905
CAGEPGM	-0.064	-0.056	-0.045	-0.060	-0.044	-0.020
ANALTOBE			-0.00562			-0.0117

The first column in Table 24 gives the variable names: in addition to an overall intercept, three district variables and six other variables are listed. The district variables are zero-one variables indicating districts A, B, and C. District H/I is the omitted district in this parametrization. District F does not participate in the propensity score modeling because it has no one-program immersion strategy school; district F students have an infinite propensity to be in the early-exit program.

The other six variables include two described earlier: EDAVG (the parents' average years of education) and PRESCHY (whether the student attended preschool). BOOKSRD gives the number of books the parent had read in the last three months (question 37b from the parent interview). RPTOCLSY has the value 1 if the parent or someone else reads to the child in Spanish and 0 otherwise (from question 43 on the parent interview). CAGEPGM is the child's age upon entering the program in kindergarten, measured in months (derived from the student database). ANALTOBE is the sum of the language and mathematics subtests from the fall kindergarten administration of the TOBE.

The PONEK1 column shows the model for the K-1 analyses using all students (PONE stands for Propensity in ONE-program schools). PONEK1PN gives the same model estimated using only students with pretest scores available. PONEK1PP includes the students with pretest scores available and adds the pretest total (ANALTOBE) as a predictor of propensity. The PONE13, PONE13PN, and PONE13PP columns give the corresponding models for the 1-3 analyses.

The coefficients in each column are very similar, indicating that all six propensity models are essentially the same. The propensity scores estimated from each of the six models were correlated. The number of students correlated ranged from 274 (the number of students in the "students with pretest" 1-3 analyses) to 423 (the number of students in the "all students" 1-3 analyses). The correlations among the six versions of the propensity score were extremely high, ranging from .8875 to .9924.

These high correlations indicate that the propensity model is insensitive to the reduction of the sample to just those students with pretest or to just those students in the 1-3 analyses. In addition, it indicates that the propensity score is little changed by the addition of the pretest as a predictor. This implies that, for the one-program schools, the students in the two programs have similar pretest scores after accounting for district membership and the other five variables.

What SES differences were found among the late-exit districts?

As was done for the IS/EE districts, school-level measures of socioeconomic status of the students were compared across the late-exit districts. The only study school in district D was assumed to represent that district, and the school-to-school variation within districts was estimated using only districts E and G. Both unweighted analyses and analyses weighted by the number of study students in the school were performed. Table 25 shows the results of those comparisons of the school-level SES measures across the three late-exit districts. None of the comparisons reach statistical significance at the .05 level, and only the unweighted analysis for the percent low income achieves significance at the .10 level. This means that no district has substantially higher or lower SES at the school level.

Table 25

Late-Exit District Differences on School SES Variables
Unweighted and Weighted by Number of Study Students

Variable	Unadjusted		Adjusted for Program	
	F-Stat	p-Value	F-Stat	p-Value
% low income	3.30	.0900	2.25	.1678
% low-middle	0.72	.5179	1.29	.3267
% middle & up	2.78	.1212	2.46	.1471
% unemployed	0.20	.8236	0.52	.6112
% unskilled labor	2.05	.1909	0.95	.4264
% blue collar	0.66	.5412	1.27	.3316
% clerical, white collar, professional	2.02	.1952	1.84	.2196
% free or reduced price meals	0.89	.4489	0.79	.4869
Overall MANOVA (Wilks' Criterion) (exact d.f.)	0.51 (16, 2)	.8279	0.96 (16, 2)	.6251

How well did the propensity models fit the data?

Generally the propensity models provided an adequate fit to the data. This was confirmed by inspecting the fitted values and residuals from the model in various ways. The residuals were checked for approximate normality and for both univariate and multivariate outliers. The residuals were also correlated with variables not included in the model to check other possible predictors and to check for attrition effects. The fitted values were checked to ensure that the values were reasonable, including checks for both univariate and multivariate outliers. These checks on the fitted values and the residuals sometimes revealed that a model was unstable, usually because a fitted parameter was based on information from too few students; those unstable models are not presented here.

Generalizability

Are findings from this study applicable to all English immersion strategy, early-exit, and late-exit instructional programs?

No. First, study results are relevant only to those programs serving Spanish-speaking language-minority students. Research suggests that second language learners of English with a primary language other than Spanish acquire English language skills differently.

Secondly, study results are applicable only to those instructional programs exhibiting the same characteristics as those in this study. The research objective was to examine three specific instructional treatments. Structured English immersion strategy and late-exit programs which most resembled each instructional model were selected from the field in 1983. In effect, these programs represented the optimal (and not the range of) implementation of each instructional model. Further, because of the need in the research design to maximize the comparisons between the instructional models, early-exit programs were selected from school districts that also had a structured English immersion strategy program. Thus, they also are not representative of all early-exit programs.

In sum, study findings are limited to programs serving Spanish-speaking language-minority students and that exhibit the same characteristics as the study programs selected.

III. THE RELATIVE EFFECTIVENESS OF IMMERSION STRATEGY AND EARLY-EXIT TRANSITIONAL BILINGUAL EDUCATION PROGRAMS IN TWO-PROGRAM SCHOOLS

Introduction

What was the purpose of the two-program schools analyses?

The analyses of two-program schools were designed to determine the relative effectiveness of the immersion strategy (IS) and early-exit (EE) programs, controlling for school effects. Program effectiveness was evaluated in three areas of academic achievement, as measured in English: mathematics, language arts, and reading.

Why were the two-program schools analyzed separately?

The four schools with both IS and EE programs provided an opportunity to evaluate program effects while controlling all factors associated with the school level and the district level. When a single school has both programs, the IS and EE programs can be compared without concern that differences between schools or districts are affecting the results. In addition, controlling for school effectively controls for the attributes of the community served by the school.

Were the IS and EE students comparable in the two-program schools at the beginning of the study?

An exhaustive examination of the students in the two-program schools, after adjusting for selection, failed to uncover any statistically significant differences between IS and EE students other than those directly related to the program. That is, there were no detectable "selection effects" in the two-program schools. Upon entering their program in kindergarten, the IS students in a school were not different in any material way in background or initial achievement from the EE students in the same school.

ERIC

The absence of any selection effects reflects a strength of the analyses focusing on two-program schools. Although students were not assigned to the IS and EE programs according to a formal randomization mechanism, as they would be in a randomized experiment, in these two-program schools the children in the two programs do not differ systematically. This means that it is reasonable to analyze the two programs as though they were from a randomized experiment.

What statistical methods were used to compare the two programs?

The statistical analyses used to evaluate the programs were performed in two parts. This was necessary because no CTBS test scores in English were available for kindergarten, making individual growth curve modeling for kindergarten through third grade inappropriate. Fall kindergarten test scores were from the TOBE test administered in the children's dominant language, almost always Spanish.

The first set of analyses were carried out in order to evaluate the achievement of each child at the spring of first grade using background information and, if available, fall kindergarten test scores as a "pretest." These analyses, called the "K-1" analyses, were performed separately for the mathematics, language arts, and reading subtests. Analysis of Covariance (ANCOVA) was used for these analyses.

The second set of analyses were done to evaluate growth from first through third grade based on the calculation of an individual growth curve for each child. These "1-3" analyses estimated growth curves extending from the spring of first grade through the spring of third grade (or through second grade if third grade test scores were unavailable). (See Appendix C for unadjusted and adjusted test scores, by program and grade.) For each of the three outcome measures, these individual growth curves were then related to background information associated with the child as well as the school and program. Specifically, a computer program was used to develop a hierarchical linear model (HLM), as described in more detail in Chapter II.

ERIC

What form of the CTBS achievement subtest scores were analyzed?

The achievement subtest scores for the spring of first grade through the spring of third grade were in the form of an Expanded Scale Score (ESS) from the CTBS test, administered in English. The ESS form a continuous scale that covers the range of achievement from kindergarten through high school. The ESS are designed so that a difference of one point has the same meaning regardless of the magnitude of the scores.

What measure of time was used in the evaluation of growth?

To evaluate growth, a measure of time is needed. The date of each administration of the achievement tests was known for each student. These dates formed the basis for the time metric. Time was measured in years of schooling, with the "zero" point taken to be spring of first grade. The advantage of setting the "zero" to spring of first grade is that the estimated "intercept" for the 1-3 growth curve analyses is the same as the end point of the K-1 analyses. Spring of second grade corresponds to time one, while spring of third grade corresponds to time two.

Were the schools significantly different from each other?

Even after including the background variables in the model, statistically significant school effects were found. School was found to be a significant predictor of both initial status and growth rate. However, it should be kept in mind that the extent to which these school differences can be attributed to measured variation in the program implementation of the schools is beyond the scope of nominal program analyses.

What are the implications of school differences?

The presence of school differences does not interfere with the comparison of the immersion strategy and early-exit programs in the

ERIC

two-program schools. The comparison of the programs controls for the school differences.

In contrast, a statistical interaction between program and school in the two-program schools would complicate the comparison between the immersion strategy and early-exit programs. Such an interaction would mean that the difference between the achievement of early-exit students and the achievement of immersion strategy students would be different in the different schools. This could mean that in some schools one program was superior but in other schools the other program was superior. Statistical tests for the presence of school-program interactions in the two-program schools were not significant in either the K-1 or the 1-3 analyses. That is, differences in program effects did not differ significantly from school to school. Thus, any difference between the immersion strategy and early-exit programs does not vary significantly across the four two-program schools.

<u>How were program differences in spring of first grade evaluated?</u>

Program differences in spring of first grade were evaluated using Analysis of Covariance (ANCOVA), controlling for school and predictors of achievement (covariates, such as nominal program label, school site, absences per year, average educational level of parents, preschool experience, gender of student, number of books in the home, TOBE math pretest, and the sum of TOBE math and language pretest scores). For those students where the fall kindergarten TOBE scores were available, these were used as a "pretest." Because the IS and EE students in each school had similar values on the predictors of achievement, the estimated difference between the programs is little affected by the inclusion of these covariates. However, the covariates improve the accuracy of the prediction of achievement at spring of first grade, permitting relatively precise estimation of school and program differences. That is, using the covariates to predict achievement increases the statistical power of the ANCOVA analyses, reducing the estimated standard error and making even relatively small differences statistically significant.

ERIC

Results of the K-1 Analyses

What were the results of the K-1 analyses for the mathematics subtest?

An extensive series of ANCOVA analyses were used to evaluate the spring first grade difference between immersion strategy and early-exit students on the mathematics subtest. Five background variables were found to have predictive power in the two-program schools analyses: EDAVG (the average years of education of the student's parents), PRESCHY (whether the student attended preschool), FEMALE (whether the student was a girl), BOOKSHM (the number of books in the student's home), and ANYEPTOP (whether the parents reported using English when talking to each other). The program difference was evaluated after controlling for all five of these background variables, the number of absences in first grade (ABS1), and the school. For comparison, the program difference was also evaluated without controlling for any variables; controlling only for school; and controlling for school, absences, and preschool attendance. (Preschool attendance was the only statistically significant background variable.) The tables summarizing the results of these analyses for mathematics will be described in considerable detail, but the discussion of tables for language and reading will be limited to pointing out highlights.

Table 26 summarizes the results of the analysis of program differences not controlling for any variable other than program. The first column, labeled "Predicted," contains the phrase "Spring 1st" as a reminder that these K-1 analyses are predicting the CTBS subtest score in spring of first grade. The second column, "Predictor," indicates the variable used as a predictor. The variable BASE represents an overall average. (Although this variable is sometimes termed the intercept, that term will be avoided here to prevent confusion with the intercept of the growth curve in the 1-3 analyses.) The variable PGMIS is a binary, or zero-one, variable that takes the value one for students in the immersion strategy program and the value zero for students in the early-exit program. The columns labeled "Parameter" and "Std Err" contain the estimated coefficient of the variable and its estimated standard error. The last two columns contain a t-statistic formed by dividing the parameter by its standard error (t-Stat) and the two-tailed probability associated with the t-statistic (p-Value).

Table 26

Two-Program Schools K-1 Analysis for Mathematics
All Students; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	247.364	3.051	81.066	0.000
Spring 1st	PGMIS	-3.951	3.799	-1.040	0.299

In the simple model represented by Table 26, each student's predicted achievement depends only on program. In general, the predicted achievement for any student is calculated by multiplying that student's values for each variable by the parameter estimate. The BASE variable has the value 1 for all students, so the coefficient for BASE is always included. A student in the early-exit program in one of these four two-program schools has the value zero for PGMIS, and so has a predicted achievement of about 247. A student in the immersion strategy program has a predicted achievement nearly four points lower, about 243 (247.364 - 3.951 = 243.413). Thus the coefficient of PGMIS can be interpreted as the difference between the immersion strategy students and the early-exit students, with a positive sign indicating IS students have higher achievement and a negative sign indicating that IS students have lower achievement.

Because the p-value associated with the PGMIS variable is about .3, a difference as large as this four-point difference between programs can be expected to occur by chance about thirty percent of the time. The most common criterion for declaring a difference to be statistically significant is that it would occur by chance less than five percent of the time (a p-value of .05 or less). The p-value of .3 does not approach this level, so the difference between programs in this analysis is not statistically significant. The "null hypothesis" that the students in the two programs do not differ in mathematics achievement in spring of first grade is not rejected.

Table 27 summarizes the results of the analysis of program differences after controlling for school. Three variables are added to allow for differences among the four two-program schools. A student in school 12 has the value 1 for SCHOOL12 and values of 0 for SCHOOL14 and SCHOOL51. A student in school 14 has the value 1 for SCHOOL14 and values of 0 for SCHOOL12 and SCHOOL51, and similarly for a student in school 51. A student in the "omitted school," school 21, has the value -1 for SCHOOL12, SCHOOL14, and SCHOOL51. One school must be omitted in any analysis of this type for statistical reasons, but the choice of which school to omit does not affect the statistical results.

Table 27

Two-Program Schools K-1 Analysis for Mathematics
All Students; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	249.236	3.074	81.069	0.000
Spring 1st	PGMIS	-6.047	3.737	-1.618	0.107
Spring 1st	SCHOOL12	-0.471	2.727	-0.173	0.863
Spring 1st	SCHOOL14	-6.811	3.155	-2.159	0.032
Spring 1st	SCHOOL51	15.622	3.697	4.226	0.000

While there are many other ways to code school variables that produce the same model, the coding used here has certain advantages. One advantage is that the coefficient for BASE is interpreted as the estimated achievement for early-exit students averaged over the four schools. Also, as it stands the coefficient can be meaningfully interpreted without additional calculations as the average achievement directly. Another advantage is that the coefficients for the SCHOOL variables are interpreted as the difference between achievement levels in that school and the average of all four schools. For example, the parameter estimate for the variable SCHOOL12 is -0.471, which is not statistically significant (p = .863). (Note that the mean difference between any two schools can be obtained by subtracting their parameters.) This means that the average achievement in school 12 is about the same as the average achievement across all four schools. In contrast, school 14 has achievement nearly seven points lower than the average across the four schools (difference = -6.811), a difference that is statistically significant (p = .032). Finally, SCHOOL51 has the value 15.622, indicating that school 51 has achievement much higher than the average of the four schools. This difference is highly statistically significant, with a p-value less than .0005 (indicated by .000 in the table).

No coefficient is given for school 21 in Table 27 but the difference between school 21 and the average of all four schools can be calculated from the coefficients given. Since students in school 21 have a value of -1 for all three SCHOOL variables, the estimated difference between school 21 and the four-school average can be calculated as the negative of the sum of the three coefficients in the table. School 21 is estimated to have achievement 8.340 points below the average of the four schools: (-1) x (-0.471) + (-1) x (-6.811) + (-1) x (15.622) = -8.340. Although the tables do not present a test of the statistical significance of this difference, hypothesis testing methods can be used to find that the difference is just significant (p = .041). Although the difference between school 21 and the average of the other four schools was tested in this way for every analysis, the results of the test will be given only when it is noteworthy.

Hypothesis testing can also indicate whether there are overall differences among the four schools. This test is performed by comparing the school-to-school variation among the four schools to the student-to-student variation within the schools. If the school-to-school variation is substantially larger than the student-to-student variation, the schools can be said to differ significantly. In Table 27 the school variation is highly significant (p less than .0005). Note that this test for overall school variation may be significant even though none of the individual schools differs significantly from the average. Also note that the test for overall school variation may be nonsignificant even though one or more of the tests for the individual schools is significant. A conservative approach is to declare individual schools significantly different from the average only if both the individual test and the overall test are statistically significant.

Comparing Table 27 and Table 26, the addition of the school variables has increased the size of the program effect from 3.951 to 6.047 and has slightly reduced the estimated standard error. One of the advantages of the two-program schools analyses is the ability to control for school, and the analysis with schools controlled produces a better estimate of the program difference. While this school-controlled analysis shows the immersion strategy students are about six points behind the early-exit students in each school, the difference is not statistically significant (p = .107).

Table 28 summarizes results from the analysis that included both ABS1 (the number of absences in first grade) and PRESCHY (whether the student attended preschool). PRESCHY was the only one of the five covariates found to be a significant predictor of mathematics achievement. Absences have the expected negative effect (the more days absent, the lower the test score), but it is not significant (p = .325). Attending preschool raises predicted test scores by over 10 points. After controlling for these two covariates, the program difference drops to 5.201 and is not significant (p = .163). Differences among the schools are also reduced but are still significant overall (p = .016). School 21 is nearly 5 points lower than average but the difference is not statistically significant (difference = -4.968, p = .238). Since school 51 is significantly higher than the average (difference = 12.309, p = .002) and the other three schools do not differ much from the average, it is fair to conclude that the overall school differences are due primarily to school 51. That is, including ABS1 and PRESCHY as covariates has essentially eliminated achievement differences among schools 12, 14, and 21. Because the coefficient of ABS1 is not significant and the average number of absences probably does not vary greatly across the schools, it is likely that including the variable PRESCHY has produced most of this effect.

That is, much of the variability among these three schools is probably accounted for by differences in the proportion of students who attended preschool.

Table 28

Two-Program Schools K-1 Analysis for Mathematics
All Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	246.659	3.825	64.485	0.000
Spring 1st	PGMIS	-5.201	3.717	-1.399	0.163
Spring 1st	SCHOOL12	-2.578	2.807	-0.919	0.359
Spring 1st	SCHOOL14	-4.762	3.219	-1.479	0.140
Spring 1st	SCHOOL51	12.309	3.873	3.178	0.002
Spring 1st	ABS1	-0.226	0.229	-0.987	0.325
Spring 1st	PRESCHY	10.032	3.817	2.628	0.009

Table 29 summarizes the effect of adding the other four covariates to the model. The results are not materially changed. The program difference is about the same in magnitude and remains nonsignificant. The overall school differences are still significant ($p = .021$) and can be attributed to the relatively high achievement of students in school 51.

Table 29

Two-Program Schools K-1 Analysis for Mathematics
All Students; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	241.568	5.658	42.694	0.000
Spring 1st	PGMIS	-5.676	3.722	-1.525	0.129
Spring 1st	SCHOOL12	-2.883	2.803	-1.029	0.305
Spring 1st	SCHOOL14	-5.420	3.255	-1.665	0.097
Spring 1st	SCHOOL51	11.891	3.955	3.006	0.003
Spring 1st	ABS1	-0.228	0.229	-0.998	0.320
Spring 1st	EDAVG	0.776	0.680	1.141	0.255
Spring 1st	PRESCHY	8.657	3.923	2.207	0.028
Spring 1st	FEMALE	-4.333	3.548	-1.221	0.223
Spring 1st	BOOKSHM	1.814	1.350	1.344	0.180
Spring 1st	ANYEPTOP	-2.073	5.769	-0.359	0.720

Summary

Although the immersion strategy students tend to score five or six points lower than early-exit students on the mathematics subtest, the difference is not statistically significant. Students in school 51 have significantly higher scores than students in the other three schools, whose scores are comparable.

What were the results of the K-1 analyses for the language subtest?

Table 30 shows the effect of program for the language subtest without controlling for school or any covariates. Table 31 shows the results after controlling for school. In both analyses the immersion strategy students have slightly higher scores than early-exit students (difference = 5.185 not controlling for school, difference = 4.019 controlling for school) but the difference is not statistically significant (p = .274 and p = .388, respectively). As for mathematics, school 51 has higher achievement than average (difference = 18.473, p less than .0005). School 21 has lower achievement than average (difference = -3.018) but this difference is not statistically significant (p = .550). Schools 12 and 14 have comparable achievement for language. Most of the school-to-school variability can be attributed to school 51.

Table 30

Two-Program Schools K-1 Analysis for Language
All Students; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	[illegible]SE	233.102	3.796	61.408	0.000
Spring 1st	[illegible]MIS	5.185	4.726	1.097	0.274

Table 31

Two-Program Schools K-1 Analysis for Language
All Students; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	236.814	3.826	61.895	0.000
Spring 1st	PGMIS	4.019	4.651	0.864	0.388
Spring 1st	SCHOOL12	-7.970	3.394	-2.348	0.020
Spring 1st	SCHOOL14	-7.485	3.926	-1.906	0.058
Spring 1st	SCHOOL51	18.473	4.600	4.015	0.000

For the language subtest, the key covariate was found to be parents' use of English as measured by ANYEPTOP. This variable takes the value 1 if the parent interview indicates that the parents use any English when speaking to each other and 0 if the parents use only Spanish. Table 32 shows the analysis of language test scores controlling for school, absences, and parental English. Parental use of English raises the estimated language test score by nearly 15 points, a significant increase (difference = 14.706, p = .036). The program effect changes only by about one point compared with the school-controlled analysis (from 4.019 to 5.004) and again it is not statistically significant (p = .282). Estimated school effects also change relatively little. Adding the other four covariates (see Table 33) has essentially no effect on the estimates of program and school effects.

Table 32

Two-Program Schools K-1 Analysis for Language
All Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	237.762	4.467	53.228	0.000
Spring 1st	PGMIS	5.004	4.638	1.079	0.282
Spring 1st	SCHOOL12	-8.204	3.396	-2.416	0.017
Spring 1st	SCHOOL14	-6.165	3.949	-1.561	0.120
Spring 1st	SCHOOL51	14.922	4.796	3.112	0.002
Spring 1st	ABS1	-0.392	0.286	-1.373	0.171
Spring 1st	ANYEPTOP	14.706	6.968	2.111	0.036

Table 33

Two-Program Schools K-1 Analysis for Language
All Students; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	231.664	7.131	32.485	0.000
Spring 1st	PGMIS	5.222	4.691	1.113	0.267
Spring 1st	SCHOOL12	-8.830	3.533	-2.499	0.013
Spring 1st	SCHOOL14	-6.339	4.103	-1.545	0.124
Spring 1st	SCHOOL51	14.277	4.985	2.864	0.005
Spring 1st	ABS1	-0.398	0.288	-1.381	0.169
Spring 1st	EDAVG	0.432	0.858	0.503	0.615
Spring 1st	PRESCHY	1.701	4.945	0.344	0.731
Spring 1st	FEMALE	1.029	4.471	0.230	0.818
Spring 1st	BOOKSHM	1.342	1.702	0.788	0.431
Spring 1st	ANYEPTOP	12.506	7.271	1.720	0.087

Summary

Although the immersion strategy students tend to score four or five points higher than early-exit students on the language subtest, the difference is not statistically significant. Students in school 51 have significantly higher scores than students in the other three schools. School 21 has approximately average scores while schools 12 and 14 have somewhat lower scores.

ERIC

What were the results of the K-1 analyses for the English reading subtest?

Table 34 presents the results of the English reading subtest analysis without controlling for school or any covariates, while Table 35 presents the school-controlled results. In both analyses the immersion strategy students have lower scores than early-exit students (difference = -7.147 not controlling for school, difference = -8.977 controlling for school). The difference is not quite statistically significant without controlling for school (p = .065) but is significant in the analysis with school controlled (p = .019). As for mathematics and reading, school 51 has higher achievement than average (difference = 14.291, p less than .0005). School 14 has significantly lower achievement than average (difference = -10.736, p = .001). Schools 12 and 21 have comparable achievement for reading, both near the average for the four schools (difference = -2.753 for school 12, difference = -0.802 for school 21). Most of the school-to--school variability can be attributed to schools 14 and 51.

Table 34

Two-Program Schools K-1 Analysis for Reading
All Students; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	227.659	3.100	73.446	0.000
Spring 1st	PGMIS	-7.147	3.859	-1.852	0.065

Table 35

Two-Program Schools K-1 Analysis for Reading
All Students; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	230.693	3.120	73.940	0.000
Spring 1st	PGMIS	-8.977	3.792	-2.367	0.019
Spring 1st	SCHOOL12	-2.753	2.767	-0.995	0.321
Spring 1st	SCHOOL14	-10.736	3.202	-3.353	0.001
Spring 1st	SCHOOL51	14.291	3.751	3.809	0.000

For the reading subtest, as for the language subtest, the key covariate was found to be ANYEPTOP: whether the parents used English with each other. Table 36 shows the analysis of reading test scores controlling for school, absences, and parental English. Parental use of English raises the estimated reading test score by nearly 14 points, a significant increase (difference = 13.951, p = .015). The program effect changes by less than one point compared with the school-controlled analysis (from -8.977 to -8.177) and it is still statistically significant (p = .031). Estimated school effects change relatively little. The inclusion of the other four covariates (Table 37) has a negligible effect on the estimated program effect. The school-to-school variability is somewhat reduced but is still statistically significant (p = .014).

Table 36

Two-Program Schools K-1 Analysis for Reading
All Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	230.888	3.634	63.537	0.000
Spring 1st	PGMIS	-8.177	3.773	-2.167	0.031
Spring 1st	SCHOOL12	-2.839	2.763	-1.027	0.305
Spring 1st	SCHOOL14	-9.671	3.213	-3.010	0.003
Spring 1st	SCHOOL51	11.224	3.901	2.877	0.004
Spring 1st	ABS1	-0.281	0.232	-1.209	0.228
Spring 1st	ANYEPTOP	13.951	5.668	2.461	0.015

Table 37

Two-Program Schools K-1 Analysis for Reading
All Students; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	232.272	5.766	40.284	0.000
Spring 1st	PGMIS	-8.433	3.793	-2.223	0.027
Spring 1st	SCHOOL12	-3.700	2.856	-1.295	0.197
Spring 1st	SCHOOL14	-8.528	3.317	-2.571	0.011
Spring 1st	SCHOOL51	9.596	4.031	2.381	0.018
Spring 1st	ABS1	-0.315	0.233	-1.351	0.178
Spring 1st	EDAVG	-0.709	0.693	-1.022	0.308
Spring 1st	PRESCHY	5.001	3.998	1.251	0.212
Spring 1st	FEMALE	-3.251	3.615	-0.899	0.369
Spring 1st	BOOKSHM	1.839	1.376	1.337	0.183
Spring 1st	ANYEPTOP	13.765	5.879	2.341	0.020

Summary

The immersion strategy students tend to score eight to nine points lower than early-exit students on the reading subtest, and the difference is statistically significant at the .05 level. Students in school 51 have significantly higher scores and students in school 14 significantly lower scores than students in the other two schools.

Including the Pretest as a Predictor in the K-1 Analyses

How were the K-1 analyses for mathematics affected by controlling for pretest?

Although the K-1 analyses estimate program effects after controlling for school and several covariates, it is possible that the results would be changed by controlling for pretest. As noted above, the pretest scores are fall kindergarten TOBE scores for mathematics and language measured in the student's dominant language, which was almost always Spanish. These pretest scores are unavailable for many students because of the design of the study. Among the four two-program schools, a majority of the students

ERIC

in schools 12 and 14 have pretest scores and all of the students in school 21 do. However, none of the students in school 51, the school with highest achievement at first grade, have pretest scores.

Before evaluating the effect of including the pretest as a covariate, it is necessary to evaluate the effect of reducing the students analyzed to those who have a pretest. This means leaving out all school 51 students. Table 38 summarizes the unadjusted program difference on the mathematics subtest for the students with a pretest and Table 39 summarizes the analysis with school differences controlled. These tables may be compared with Table 26 and Table 27, respectively.

Table 38

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	234.333	3.883	60.349	0.000
Spring 1st	PGMIS	3.990	4.687	0.851	0.396

Table 39

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	234.8[illegible]	3.950	59.455	0.000
Spring 1st	PGMIS	2.957	5.014	0.598	0.551
Spring 1st	SCHOOL12	1.431	3.121	0.458	0.647
Spring 1st	SCHOOL14	-2.550	3.357	-0.760	0.449

Among the students with a pretest, the immersion strategy students have slightly higher mathematics test scores than the early-exit students, but the difference is not statistically significant. Recall that for all students, the IS students had slightly lower scores (but also not

654

statistically significant). Separate analyses not presented here indicate that this change of sign is primarily due to the exclusion of school 51. The fact that the program effect changes sign when omitting a school implies that the program effect differs among the schools, indicating a program-by-school interaction. The program effect was not significant either in the analysis of all students or in the analysis of the students with pretest. Furthermore, explicit tests for program-by-school interaction for the mathematics subtest were not statistically significant. These findings, including the change in the sign of the estimated program effect after omitting school 51, serve to emphasize the conclusion that the IS and EE students have comparable achievement on the mathematics subtest.

Table 39 includes only two SCHOOL variables, since school 51 cannot be included in the analyses of students with pretest. The test for school differences is not significant (p = .725), confirming the finding from the all-students analyses that school differences among schools 12, 14, and 21 are relatively small in mathematics. Table 40 which corresponds to Table 28 and Table 41 which corresponds to Table 29 present the covariate-adjusted analyses for students with pretest. The program effect continues to be positive and not near statistical significance, and the test for overall school differences also is not statistically significant. The effects of the covariates are similar in the two sets of analyses.

Table 40

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	233.631	4.763	49.056	0.000
Spring 1st	PGMIS	3.899	4.992	0.781	0.436
Spring 1st	SCHOOL12	-2.026	3.459	-0.586	0.559
Spring 1st	SCHOOL14	-1.127	3.384	-0.333	0.740
Spring 1st	ABS1	-0.241	0.279	-0.863	0.390
Spring 1st	PRESCHY	10.709	5.028	2.130	0.035

Table 41

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	226.039	7.415	30.486	0.000
Spring 1st	PGMIS	2.500	5.112	0.489	0.626
Spring 1st	SCHOOL12	-2.000	3.477	-0.575	0.566
Spring 1st	SCHOOL14	-2.652	3.540	-0.749	0.455
Spring 1st	ABS1	-0.255	0.281	-0.908	0.365
Spring 1st	EDAVG	1.502	0.924	1.626	0.106
Spring 1st	PRESCHY	6.977	5.525	1.263	0.209
Spring 1st	FEMALE	-2.749	4.535	-0.606	0.545
Spring 1st	BOOKSHM	0.793	1.770	0.448	0.655
Spring 1st	ANYEPTOP	2.234	8.600	0.260	0.795

Table 42, Table 43, Table 44, and Table 45 show the results of including both the mathematics pretest (ANALPREM) and the language pretest (ANALPREL) as predictors of mathematics achievement in spring of first grade. While the mathematics pretest is always a highly significant predictor (p less than .0005), the language pretest is not even close to significance as an additional covariate. Including the pretests as a predictor increases the estimated program effect by about two points (from a range of 2.5 to 4.0 to a range of 4.5 to 6) but the difference is never statistically significant. School differences also remain nonsignificant.

Table 42

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	130.535	25.080	5.205	0.000
Spring 1st	PGMIS	5.848	4.485	1.304	0.194
Spring 1st	ANALPREM	0.589	0.162	3.627	0.000
Spring 1st	ANALPREL	0.046	0.089	0.519	0.605

Table 43

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, School, and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	124.836	25.471	4.901	0.000
Spring 1st	PGMIS	5.086	4.777	1.065	0.289
Spring 1st	ANALPREM	0.638	0.165	3.855	0.000
Spring 1st	ANALPREL	0.038	0.090	0.419	0.676
Spring 1st	SCHOOL12	0.719	2.985	0.241	0.810
Spring 1st	SCHOOL14	-4.792	3.236	-1.481	0.141

Table 44

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Key Covariates and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	126.826	25.347	5.004	0.000
Spring 1st	PGMIS	5.838	4.755	1.228	0.222
Spring 1st	ANALPREM	0.646	0.164	3.941	0.000
Spring 1st	ANALPREL	0.012	0.090	0.131	0.896
Spring 1st	SCHOOL12	-2.402	3.284	-0.731	0.466
Spring 1st	SCHOOL14	-3.508	3.260	-1.076	0.284
Spring 1st	ABS1	-0.238	0.265	-0.897	0.371
Spring 1st	PRESCHY	9.952	4.802	2.072	0.040

Table 45

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; All Covariates and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	126.247	25.612	4.929	0.000
Spring 1st	PGMIS	4.482	4.894	0.916	0.361
Spring 1st	ANALPREM	0.603	0.167	3.603	0.000
Spring 1st	ANALPREL	0.029	0.093	0.314	0.754
Spring 1st	SCHOOL12	-2.464	3.325	-0.741	0.460
Spring 1st	SCHOOL14	[illegible]321	3.413	-1.266	0.208
Spring 1st	ABS1	[illegible].239	0.268	-0.893	0.373
Spring 1st	EDAVG	1.006	0.889	1.131	0.260
Spring 1st	PRESCHY	7.350	5.288	1.390	0.167
Spring 1st	FEMALE	-2.995	4.422	-0.677	0.499
Spring 1st	BOOKSHM	0.856	1.686	0.508	0.612
Spring 1st	ANYEPTOP	0.599	8.203	0.073	0.942

Table 46 through Table 49 reflect the corresponding analyses with the sum of the mathematics and language pretests (ANALTOBE) as a covariate. The analysis determined in this instance that the overall TOBE score was a better predictor than the subtest scores taken individually. The sum of the mathematics and language TOBE subtest scores is the overall score on the TOBE. The estimated program effects are between those from the analyses of students who have a pretest (but not using the pretests as a predictor) and the analyses using both pretests as predictors. The program difference never approaches statistical significance and school differences are also nonsignificant.

Table 46

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	165.634	21.305	7.774	0.000
Spring 1st	PGMIS	5.138	4.557	1.128	0.261
Spring 1st	ANALTOBE	0.208	0.064	3.276	0.001

Table 47

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, School, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	163.946	21.654	7.571	0.000
Spring 1st	PGMIS	4.740	4.881	0.971	0.333
Spring 1st	ANALTOBE	0.215	0.065	3.325	0.001
Spring 1st	SCHOOL12	0.172	3.045	0.057	0.955
Spring 1st	SCHOOL14	-3.188	3.255	-0.979	0.329

Table 48

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Key Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	167.674	21.739	7.706	0.000
Spring 1st	PGMIS	5.390	4.876	1.105	0.271
Spring 1st	ANALTOBE	0.201	0.065	3.102	0.002
Spring 1st	SCHOOL12	-2.669	3.368	-0.793	0.429
Spring 1st	SCHOOL14	-1.942	3.300	-0.589	0.557
Spring 1st	ABS1	-0.206	0.272	-0.760	0.449
Spring 1st	PRESCHY	9.041	4.916	1.839	0.068

Table 49

Two-Program Schools K-1 Analysis for Mathematics
Students with Pretest: All Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	162.211	21.974	7.382	0.000
Spring 1st	PGMIS	3.651	4.982	0.733	0.465
Spring 1st	ANALITOBE	0.202	0.066	3.075	0.003
Spring 1st	SCHOOL12	-2.820	3.389	-0.832	0.407
Spring 1st	SCHOOL14	-3.000	3.442	-0.872	0.385
Spring 1st	ABS1	-0.212	0.273	-0.777	0.439
Spring 1st	EDAVG	1.275	0.901	1.415	0.159
Spring 1st	PRESCHY	6.186	5.374	1.151	0.252
Spring 1st	FEMALE	-4.776	4.455	-1.072	0.286
Spring 1st	BOOKSHM	0.777	1.720	0.452	0.652
Spring 1st	ANYEPTOP	1.086	8.365	0.130	0.897

Summary

Reducing the sample to students with pretest scores changes the direction of the program difference from favoring early-exit to favoring immersion strategy, but it is still not statistically significant. The change in direction is due primarily to the absence of school 51, the school with highest achievement. Among the students with pretest at the remaining three schools, the overall school variation is not significant. Adding the mathematics and language pretest scores as covariates -- either as separate variables or as the sum -- tends to increase the estimated superiority of IS students slightly but the difference remains nonsignificant. School differences also remain nonsignificant. In short, for the students with pretest scores, neither program nor school effects manifest themselves; it is a fair characterization to say that the two programs and all three schools have approximately the same level of mathematics achievement.

How were the K-1 analyses for language affected by controlling for pretest?

The evaluation of the effect of including the pretest as a covariate for the language subtest parallels the evaluation for the mathematics subtest. Table 50 through Table 53 summarize the estimated models when the sample is limited to the students with a pretest. Unlike the effect for mathematics, reducing to the students with pretest has a relatively small effect on the estimated program difference. The small advantage in language achievement at first grade for IS students is somewhat increased, but the difference is still not statistically significant. As for mathematics, the three remaining schools do not differ significantly.

Table 50

Two-Program Schools K-1 Analysis for Language
Students with Pretest; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	229.125	4.996	45.865	0.000
Spring 1st	PGMIS	6.675	6.030	1.107	0.270

Table 51

Two-Program Schools K-1 Analysis for Language
Students with Pretest; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	229.156	5.085	45.061	0.000
Spring 1st	PGMIS	7.596	6.456	1.177	0.241
Spring 1st	SCHOOL12	-1.996	4.019	-0.497	0.620
Spring 1st	SCHOOL14	-0.682	4.322	-0.158	0.875

ERIC

Table 52

Two-Program Schools K-1 Analysis for Language
Students with Pretest; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	231.234	5.937	38.945	0.000
Spring 1st	PGMIS	7.984	6.464	1.235	0.219
Spring 1st	SCHOOL12	-2.979	4.100	-0.727	0.469
Spring 1st	SCHOOL14	-0.910	4.320	-0.211	0.833
Spring 1st	ABS1	-0.357	0.364	-0.981	0.328
Spring 1st	ANYEPTOP	16.263	10.967	1.483	0.140

Table 53

Two-Program Schools K-1 Analysis for Language
Students with Pretest; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	228.575	9.746	23.454	0.000
Spring 1st	PGMIS	8.263	6.720	1.230	0.221
Spring 1st	SCHOOL12	-2.648	4.570	-0.579	0.563
Spring 1st	SCHOOL14	-1.437	4.654	-0.309	0.758
Spring 1st	ABS1	-0.357	0.369	-0.968	0.335
Spring 1st	EDAVG	-0.057	1.214	-0.047	0.962
Spring 1st	PRESCHY	-1.145	7.262	-0.158	0.875
Spring 1st	FEMALE	3.071	5.961	0.515	0.607
Spring 1st	BOOKSHM	0.991	2.326	0.426	0.671
Spring 1st	ANYEPTOP	16.071	11.304	1.422	0.157

Table 54, Table 55, Table 56, and Table 57 show the results of including both the mathematics pretest (ANALPREM) and the language pretest (ANALPREL) as predictors of language achievement in spring of first grade. While the language pretest is always a significant predictor (p = .010 or less), the mathematics pretest is not even close to significance as an additional covariate. Including the pretests as a predictor increases the estimated program effect by one to two points (from a range of 6.7 to 8.3 to a range of 7.6 to 9.8) but the difference is never statistically significant. School differences also remain nonsignificant.

ERIC

Table 54

Two-Program Schools K-1 Analysis for Language Students with Pretest; Program and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	167.293	33.111	5.052	0.000
Spring 1st	PGMIS	7.648	5.921	1.292	0.198
Spring 1st	ANALPREM	0.057	0.214	0.264	0.792
Spring 1st	ANALPREL	0.315	0.118	2.676	0.008

Table 55

Two-Program Schools K-1 Analysis for Language Students with Pretest; Program, School and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	161.875	33.753	4.796	0.000
Spring 1st	PGMIS	9.507	6.330	1.502	0.135
Spring 1st	ANALPREM	0.077	0.219	0.352	0.726
Spring 1st	ANALPREL	0.328	0.119	2.745	0.007
Spring 1st	SCHOOL12	-3.709	3.956	-0.938	0.350
Spring 1st	SCHOOL14	-0.765	4.289	-0.178	0.859

Table 56

Two-Program Schools K-1 Analysis for Language Students with Pretest; Key Covariates and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	166.798	33.933	4.915	0.000
Spring 1st	PGMIS	9.757	6.350	1.537	0.127
Spring 1st	ANALPREM	0.068	0.219	0.308	0.759
Spring 1st	ANALPREL	0.318	0.119	2.662	0.009
Spring 1st	SCHOOL12	-4.498	4.036	-1.115	0.267
Spring 1st	SCHOOL14	-0.941	4.292	-0.219	0.827
Spring 1st	ABS1	-0.305	0.355	-0.859	0.392
Spring 1st	ANYEPTOP	14.298	10.735	1.332	0.185

Table 57

Two-Program Schools K-1 Analysis for Language
Students with Pretest; All Covariates and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	166.642	34.522	4.827	0.000
Spring 1st	PGMIS	9.298	6.597	1.409	0.161
Spring 1st	ANALPREM	0.067	0.225	0.297	0.767
Spring 1st	ANALPREL	0.324	0.125	2.600	0.010
Spring 1st	SCHOOL12	-3.810	4.482	-0.850	0.397
Spring 1st	SCHOOL14	-1.269	4.600	-0.276	0.783
Spring 1st	ABS1	-0.293	0.361	-0.810	0.419
Spring 1st	EDAVG	-0.215	1.199	-0.179	0.858
Spring 1st	PRESCHY	-2.633	7.128	-0.369	0.712
Spring 1st	FEMALE	-0.207	5.960	-0.035	0.972
Spring 1st	BOOKSHM	0.934	2.273	0.411	0.682
Spring 1st	ANYEPTOP	14.886	11.056	1.346	0.180

Table 58 through Table 61 reflect the corresponding analyses with the sum of the mathematics and language pretests (ANALTOBE) as a covariate. The estimated program effects are very similar to the effects estimated in the analyses using both pretests as predictors. The program difference and school differences remain nonsignificant.

Table 58

Two-Program Schools K-1 Analysis for Language
Students with Pretest; Program and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	150.569	27.616	5.452	0.000
Spring 1st	PGMIS	7.987	5.906	1.352	0.178
Spring 1st	ANALTOBE	0.238	0.082	2.890	0.004

Table 59

Two-Program Schools K-1 Analysis for Language
Students with Pretest; Program, School, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	145.557	28.049	5.189	0.000
Spring 1st	PGMIS	9.651	6.323	1.526	0.129
Spring 1st	ANALTOBE	0.254	0.084	3.028	0.003
Spring 1st	SCHOOL12	-3.481	3.944	-0.883	0.379
Spring 1st	SCHOOL14	-1.435	4.216	-0.340	0.734

Table 60

Two-Program Schools K-1 Analysis for Language
Students with Pretest; Key Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	150.570	28.337	5.314	0.000
Spring 1st	PGMIS	9.920	6.342	1.564	0.120
Spring 1st	ANALTOBE	0.244	0.084	2.908	0.004
Spring 1st	SCHOOL12	-4.292	4.025	-1.066	0.288
Spring 1st	SCHOOL14	-1.601	4.221	-0.379	0.705
Spring 1st	ABS1	-0.315	0.355	-0.889	0.376
Spring 1st	ANYEPTOP	14.123	10.724	1.317	0.190

Table 61

Two-Program Schools K-1 Analysis for Language
Students with Pretest; All Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	150.510	29.011	5.188	0.000
Spring 1st	PGMIS	9.671	6.577	1.470	0.144
Spring 1st	ANALTOBE	0.247	0.087	2.848	0.005
Spring 1st	SCHOOL12	-3.650	4.474	-0.816	0.416
Spring 1st	SCHOOL14	-1.862	4.544	-0.410	0.683
Spring 1st	ABS1	-0.305	0.361	-0.845	0.399
Spring 1st	EDAVG	-0.336	1.189	-0.282	0.778
Spring 1st	PRESCHY	-2.111	7.096	-0.298	0.767
Spring 1st	FEMALE	0.592	5.882	0.101	0.920
Spring 1st	BOOKSHM	0.970	2.270	0.427	0.670
Spring 1st	ANYEPTOP	14.668	11.044	1.328	0.186

Summary

Reducing the sample to students with pretest scores tends to increase the program effect favoring immersion strategy over early-exit but it is still not statistically significant. Adding the pretest scores as covariates slightly increases the estimated program effect but it remains nonsignificant. For these students with pretest, the overall school variation among the three schools is not significant. The finding from the analyses of all students that the IS students have a small but not statistically significant advantage is confirmed by the evaluation of the pretest as a covariate.

How were the K-1 analyses for reading affected by controlling for pretest?

The analyses of the pretest effect for the reading subtest are essentially the same as for the mathematics and language subtests. Table 62 through Table 65 summarize the models estimated for the students with pretest scores. Reducing the sample to the students with pretest has the effect of increasing the estimated early-exit advantage by about six points. The program difference, which had just achieved statistical signif:ance at the .05 level, is now significant at the .01 level. As for the mathematics and language subtests, the three remaining schools do not differ significantly in reading achievement.

Table 62

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	230.000	4.060	56.646	0.000
Spring 1st	PGMIS	-13.086	4.901	-2.670	0.008

Table 63

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	230.921	4.110	56.182	0.000
Spring 1st	PGMIS	-14.932	5.218	-2.861	0.005
Spring 1st	SCHOOL12	2.632	3.248	0.810	0.419
Spring 1st	SCHOOL14	-4.867	3.494	-1.393	0.166

Table 64

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	232.593	4.806	48.397	0.000
Spring 1st	PGMIS	-14.619	5.232	-2.794	0.006
Spring 1st	SCHOOL12	1.872	3.319	0.564	0.574
Spring 1st	SCHOOL14	-5.025	3.497	-1.437	0.153
Spring 1st	ABS1	-0.279	0.294	-0.947	0.345
Spring 1st	ANYEPTOP	11.875	8.877	1.338	0.183

Table 65

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	233.180	7.754	30.071	0.000
Spring 1st	PGMIS	-16.040	5.347	-3.000	0.003
Spring 1st	SCHOOL12	0.081	3.636	0.022	0.982
Spring 1st	SCHOOL14	-3.988	3.703	-1.077	0.283
Spring 1st	ABS1	-0.285	0.294	-0.969	0.334
Spring 1st	EDAVG	-0.713	0.966	-0.738	0.462
Spring 1st	PRESCHY	6.049	5.778	1.047	0.297
Spring 1st	FEMALE	-3.397	4.743	-0.716	0.475
Spring 1st	BOOKSHM	3.182	1.851	1.719	0.088
Spring 1st	ANYEPTOP	9.720	8.994	1.081	0.282

Including both the mathematics and language pretests (ANALPREM and ANALPREL) as covariates reduces the EE advantage by about one point, but it is still statistically significant at the .01 level (see Table 66 through Table 69). While the language pretest is always the stronger predictor it is usually nonsignificant. The p-values range from .037 (with only program and the pretests as predictors) to .062 (with the key covariates). The mathematics pretest is never close to significance. School differences remain nonsignificant. The analyses using the sum of the mathematics and language pretests (ANALTOBE) are remarkably similar (see Table 70 through Table 73). The estimated program effects scarcely change at all.

Table 66

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	202.214	27.288	7.410	0.000
Spring 1st	PGMIS	-12.677	4.880	-2.598	0.010
Spring 1st	ANALPREM	-0.038	0.177	-0.218	0.828
Spring 1st	ANALPREL	0.204	0.097	2.103	0.037

ERIC

Table 67

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, School, and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	197.638	27.755	7.121	0.000
Spring 1st	PGMIS	-13.924	5.205	-2.675	0.008
Spring 1st	ANALPREM	0.007	0.180	0.039	0.969
Spring 1st	ANALPREL	0.192	0.098	1.959	0.052
Spring 1st	SCHOOL12	1.658	3.253	0.510	0.611
Spring 1st	SCHOOL14	-4.780	3.527	-1.355	0.177

Table 68

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Key Covariates and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	201.536	27.927	7.216	0.000
Spring 1st	PGMIS	-13.707	5.226	-2.623	0.010
Spring 1st	ANALPREM	0.000002	0.181	0.000	1.000
Spring 1st	ANALPREL	0.185	0.098	1.879	0.062
Spring 1st	SCHOOL12	1.020	3.321	0.307	0.759
Spring 1st	SCHOOL14	-4.906	3.532	-1.389	0.167
Spring 1st	ABS1	-0.249	0.292	-0.852	0.396
Spring 1st	ANYEPTOP	10.863	8.835	1.230	0.221

ERIC

Table 69

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; All Covariates and Both Pretests

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	202.612	27.872	7.269	0.000
Spring 1st	PGMIS	-15.552	5.326	-2.920	0.004
Spring 1st	ANALPREM	-0.005	0.182	-0.026	0.979
Spring 1st	ANALPREL	0.198	0.101	1.964	0.052
Spring 1st	SCHOOL12	-0.601	3.619	-0.166	0.868
Spring 1st	SCHOOL14	-3.757	3.714	-1.012	0.314
Spring 1st	ABS1	-0.246	0.292	-0.844	0.400
Spring 1st	EDAVG	-0.772	0.968	-0.798	0.426
Spring 1st	PRESCHY	5.103	5.755	0.887	0.377
Spring 1st	FEMALE	-5.400	4.812	-1.122	0.264
Spring 1st	BOOKSHM	3.142	1.835	1.712	0.089
Spring 1st	ANYEPTOP	9.114	8.927	1.021	0.309

Table 70

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	186.523	22.778	8.189	0.000
Spring 1st	PGMIS	-12.360	4.871	-2.537	0.012
Spring 1st	ANALTOBE	0.132	0.068	1.939	0.054

Table 71

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, School, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	185.573	23.053	8.050	0.000
Spring 1st	PGMIS	-13.817	5.197	-2.659	0.009
Spring 1st	ANALTOBE	0.138	0.069	1.999	0.048
Spring 1st	SCHOOL12	1.826	3.241	0.563	0.574
Spring 1st	SCHOOL14	-5.275	3.465	-1.522	0.130

Table 72

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; Key Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	189.569	23.309	8.133	0.000
Spring 1st	PGMIS	-13.586	5.216	-2.605	0.010
Spring 1st	ANALTOBE	0.130	0.069	1.886	0.061
Spring 1st	SCHOOL12	1.172	3.311	0.354	0.724
Spring 1st	SCHOOL14	-5.394	3.472	-1.553	0.123
Spring 1st	ABS1	-0.257	0.292	-0.879	0.381
Spring 1st	ANYEPTOP	10.734	8.821	1.217	0.226

Table 73

Two-Program Schools K-1 Analysis for Reading
Students with Pretest; All Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	189.905	23.420	8.109	0.000
Spring 1st	PGMIS	-15.259	5.310	-2.874	0.005
Spring 1st	ANALTOBE	0.137	0.070	1.956	0.052
Spring 1st	SCHOOL12	-0.475	3.612	-0.131	0.896
Spring 1st	SCHOOL14	-4.224	3.669	-1.151	0.252
Spring 1st	ABS1	-0.256	0.291	-0.878	0.381
Spring 1st	EDAVG	-0.867	0.960	-0.903	0.368
Spring 1st	PRESCHY	5.514	5.728	0.963	0.337
Spring 1st	FEMALE	-4.771	4.749	-1.005	0.317
Spring 1st	BOOKSHM	3.171	1.833	1.730	0.086
Spring 1st	ANYEPTOP	8.942	8.915	1.003	0.318

Summary

Reducing the sample to students with pretest scores tends to increase the program effect favoring early-exit students over immersion strategy students for English reading and increases the statistical significance. This is consistent with the primary language hypothesis that teaching students to read in their first language facilitates their learning to read in a second language. The data suggest that early-exit students may be transferring their pre-reading skills from Spanish to English reading. Adding the pretest scores as covariates very slightly decreases the estimated program effect but it remains significant at the .01 level. The overall school variation for the students with a pretest is not significant. The finding from the analyses of all students that at spring of first grade the EE students have a statistically significant advantage over the IS students in English reading is confirmed by the analyses using the pretest as a covariate.

Limiting the K-1 Analyses to Students Also in the 1-3 Analyses

How were the K-1 analyses for mathematics, language, and reading affected by including only students who were also in the 1-3 analyses?

Any student with spring first grade test scores was included in the main K-1 analyses. Only students with first grade test scores and scores for at least one of either second grade or third grade were included in the 1-3 analyses. Thus the students in the 1-3 analyses are a subset of the students in the main K-1 analyses. Before the results of the K-1 analyses can be compared with the results of the 1-3 analyses, it is important to evaluate how the K-1 analyses are changed by reducing the sample to include only those students also included in the 1-3 analyses.

Table 74 shows the unadjusted program effect for mathematics, and Table 75 shows the program effect after adjusting for school. These may be compared with Table 26 and Table 27, respectively. Although the unadjusted program effect is increased by about 2.5 points, from about a 4 point advantage for early-exit to about a 6.5 point advantage for early-exit, the program effect adjusted for school increases only about 1.6 points (from 6 to 7.6 points). Although the difference between the programs is still not significant at the .05 level, after adjusting for school the difference is close to significance ($p = .063$). The overall school effect is significant ($p = .004$), with school 51 significantly higher than average (difference = 13.920, p less than .0005) and schools

12 and 14 about average. The omitted school, school 21, is lower than average (difference = -8.892) but the difference is not quite significant (p = .073).

Table 74

Two-Program Schools K-1 Analysis for Mathematics
Students in 1-3 Analysis; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	251.925	3.396	74.188	0.000
Spring 1st	PGMIS	-6.458	4.134	-1.562	0.120

Table 75

Two-Program Schools K-1 Analysis for Mathematics
Students in 1-3 Analysis; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	252.235	3.492	72.239	0.000
Spring 1st	PGMIS	-7.627	4.077	-1.870	0.063
Spring 1st	SCHOOL12	-0.204	2.966	-0.069	0.945
Spring 1st	SCHOOL14	-4.824	3.537	-1.364	0.174
Spring 1st	SCHOOL51	13.920	3.837	3.628	0.000

Table 76 shows the results for mathematics after adjusting for school, absences, and preschool attendance. The estimated difference between programs is reduced by about one point and is not significant (p = .100). The school differences are considerably reduced and are not quite significant overall (p = .051), with school 51 still significantly high (difference = 11.034, p = .007). However, school 21 is now about average (difference = -4.968, p = .238). Table 77 shows the results after adjusting for all the covariates. The program difference is increased and comes close to statistical significance (p = .056). The school differences are further reduced but have the same pattern as in the analysis adjusting for absence and preschool: overall school difference just short of significance (p = .056) but school 51 higher than average (difference = 10.471, p = .011) and the other three schools about equal to each other.

Table 76

Two-Program Schools K-1 Analysis for Mathematics
Students in 1-3 Analysis; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	249.618	4.378	57.016	0.000
Spring 1st	PGMIS	-6.779	4.099	-1.654	0.100
Spring 1st	SCHOOL12	-2.312	3.088	-0.749	0.455
Spring 1st	SCHOOL14	-3.147	3.625	-0.868	0.386
Spring 1st	SCHOOL51	11.034	4.045	2.728	0.007
Spring 1st	ABS1	-0.186	0.305	-0.610	0.542
Spring 1st	PRESCHY	8.900	4.121	2.160	0.032

Table 77

Two-Program Schools K-1 Analysis for Mathematics
Students in 1-3 Analysis; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	244.757	6.147	39.820	0.000
Spring 1st	PGMIS	-7.855	4.084	-1.923	0.056
Spring 1st	SCHOOL12	-3.304	3.074	-1.075	0.284
Spring 1st	SCHOOL14	-4.109	3.655	-1.124	0.262
Spring 1st	SCHOOL51	10.471	4.097	2.556	0.011
Spring 1st	ABS1	-0.098	0.304	-0.323	0.747
Spring 1st	EDAVG	0.747	0.747	1.000	0.318
Spring 1st	PRESCHY	8.303	4.194	1.980	0.049
Spring 1st	FEMALE	-7.087	3.847	-1.842	0.067
Spring 1st	BOOKSHM	2.313	1.448	1.598	0.112
Spring 1st	ANYEPTOP	-3.185	6.214	-0.513	0.609

Summary

The K-1 mathematics results are little changed by the reduction to students who are also in the 1-3 analyses. There is still a small advantage for early-exit students which does not achieve statistical significance. School 51 exhibits higher achievement than the other three schools, which are about equal. It is important to keep in mind the pretest analyses that show a small, nonsignificant, advantage for immersion strategy students after omitting school 51. The near-significance of the early-exit advantage in mathematics is evidently due in large part to the influence of school 51. It is safest to conclude simply that there are no significant differences between the IS and EE students in mathematics achievement in spring of first grade.

For the language subtest, the conclusions from the K-1 analyses are not changed by the reduction to students included in the 1-3 analyses (see Table 78 through Table 81). The estimated program effect is increased by about three points, from a range of a 4.0 to 5.2 point IS advantage to a range of a 7.2 to 8.6 point advantage. The program differences are still not statistically significant. The overall school effects are significant (p less than .01), with school 51 significantly higher than average and school 21 about average. Schools 12 and 14 are lower than the average of the four two-program schools. The effect for school 12 is always significant at the .01 level. The effect for school 14 is barely significant at the .05 level for the analysis with no covariates (p = .044), but it is not significant in the analyses with covariates (p = .100 and p = .080).

Table 78

Two-Program Schools K-1 Analysis for Language
Students in 1-3 Analysis; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	231.881	4.308	53.825	0.000
Spring 1st	PGMIS	7.544	5.244	1.438	0.152

Table 79

Two-Program Schools K-1 Analysis for Language Students in 1-3 Analysis; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	236.011	4.386	53.810	0.000
Spring 1st	PGMIS	7.168	5.122	1.400	0.163
Spring 1st	SCHOOL12	-10.174	3.726	-2.730	0.007
Spring 1st	SCHOOL14	-9.025	4.443	-2.031	0.044
Spring 1st	SCHOOL51	15.527	4.820	3.222	0.002

Table 80

Two-Program Schools K-1 Analysis for Language Students in 1-3 Analysis; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	235.970	5.112	46.162	0.000
Spring 1st	PGMIS	8.432	5.156	1.635	0.104
Spring 1st	SCHOOL12	-10.577	3.722	-2.842	0.005
Spring 1st	SCHOOL14	-7.482	4.521	-1.655	0.100
Spring 1st	SCHOOL51	12.262	5.043	2.431	0.016
Spring 1st	ABS1	-0.315	0.383	-0.822	0.412
Spring 1st	ANYEPTOP	15.288	7.582	2.016	0.045

ERIC

Table 81

Two-Program Schools K-1 Analysis for Language Students in 1-3 Analysis; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	227.336	7.841	28.995	0.000
Spring 1st	PGMIS	8.589	5.210	1.649	0.101
Spring 1st	SCHOOL12	-11.467	3.921	-2.925	0.004
Spring 1st	SCHOOL14	-8.218	4.662	-1.763	0.080
Spring 1st	SCHOOL51	11.463	5.226	2.193	0.030
Spring 1st	ABS1	-0.284	0.388	-0.734	0.464
Spring 1st	EDAVG	0.537	0.953	0.563	0.574
Spring 1st	PRESCHY	1.553	5.350	0.290	0.772
Spring 1st	FEMALE	1.605	4.908	0.327	0.744
Spring 1st	BOOKSHM	2.215	1.847	1.200	0.232
Spring 1st	ANYEPTOP	12.138	7.926	1.531	0.127

Summary

The K-1 results on the reduced 1-3 sample for the language subtest tend to show a stronger immersion strategy advantage but the program difference is never statistically significant. The schools differ significantly and have the same relationships as for the K-1 analyses using all students.

Including only those students who are also in the 1-3 analyses has just enough effect on the K-1 analyses for reading to eliminate the statistically significant differences between the programs. Table 82 through Table 85 show that the estimated program effect is reduced by 2.0 to 3.4 points, from a range of a 7.1 to 9.0 point EE advantage to a range of a 4.8 to 6.6 point advantage. The program differences are no longer statistically significant ($p = .122$ to $p = .260$), partly due to the reduction in the estimated difference and partly due to the increase in estimated standard error. The overall school effects are significant with no covariates ($p = .002$) but are only barely significant with absences and parental English included ($p = .045$). There are no significant school differences when all covariates are included ($p = .093$). The pattern of school differences is the same as in the K-1 reading analyses with all students (see Table 34 through Table 37).

Table 82

Two-Program Schools K-1 Analysis for Reading
Students in 1-3 Analysis; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	225.985	3.547	63.720	0.000
Spring 1st	PGMIS	-5.165	4.317	-1.196	0.233

Table 83

Two-Program Schools K-1 Analysis for Reading
Students in 1-3 Analysis; Program and School

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	227.932	3.634	62.727	0.000
Spring 1st	PGMIS	-6.585	4.243	-1.552	0.122
Spring 1st	SCHOOL12	-2.113	3.087	-0.684	0.495
Spring 1st	SCHOOL14	-9.545	3.681	-2.593	0.010
Spring 1st	SCHOOL51	13.747	3.993	3.443	0.001

Table 84

Two-Program Schools K-1 Analysis for Reading
Students in 1-3 Analysis; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	230.029	4.196	54.819	0.000
Spring 1st	PGMIS	-4.784	4.233	-1.130	0.260
Spring 1st	SCHOOL12	-2.807	3.055	-0.919	0.359
Spring 1st	SCHOOL14	-7.397	3.712	-1.993	0.048
Spring 1st	SCHOOL51	9.939	4.140	2.401	0.017
Spring 1st	ABS1	-0.610	0.314	-1.941	0.054
Spring 1st	ANYEPTOP	13.537	6.224	2.175	0.031

ERIC

Table 85

Two-Program Schools K-1 Analysis for Reading
Students in 1-3 Analysis; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	233.006	6.372	36.565	0.000
Spring 1st	PGMIS	-5.459	4.234	-1.289	0.199
Spring 1st	SCHOOL12	-4.074	3.187	-1.278	0.203
Spring 1st	SCHOOL14	-6.149	3.789	-1.623	0.106
Spring 1st	SCHOOL51	7.958	4.247	1.874	0.063
Spring 1st	ABS1	-0.605	0.315	-1.920	0.056
Spring 1st	EDAVG	-1.092	0.775	-1.410	0.160
Spring 1st	PRESCHY	5.599	4.348	1.288	0.199
Spring 1st	FEMALE	-5.240	3.989	-1.314	0.191
Spring 1st	BOOKSHM	2.590	1.501	1.726	0.086
Spring 1st	ANYEPTOP	13.945	6.442	2.165	0.032

Summary

The K-1 results for the English reading subtest for students included in the 1-3 analyses show a weaker early-exit advantage and the program difference is no longer statistically significant. The school differences are less significant but have the same relationships as for the K-1 analyses using all students.

Conclusions From the K-1 Analyses

What were the conclusions from the K-1 analyses for the mathematics subtest?

For the mathematics subtest, the K-1 analyses showed no statistically significant differences between the immersion strategy and early-exit students in the four two-program schools. The estimated program difference for each of the K-1 analyses is summarized in Table 86. The analyses of all students showed a small advantage for early-exit students. When the K-1 analyses are restricted to only those students who are also in the 1-3 analyses, a slightly larger early-exit advantage is noted, but

ERIC

the difference did not quite achieve statistical significance. The analyses controlling for school, absences, and preschool attendance showed an early-exit advantage of about 5.2 points (p = .163) for all students and about 6.8 points (p = .100) for students in the 1-3 analyses. Neither difference is near statistical significance.

Table 86

Two-Program Schools K-1 Analysis:
Estimated Program Differences in Spring 1st for Mathematics

Students; Analysis	Parameter	Std Err	t-Stat	p-Value
All Students; Program Only	-3.951	3.799	-1.040	0.299
All Students; Program & School	-6.047	3.737	-1.618	0.107
All Students; Key Covariates	-5.201	3.717	-1.399	0.163
All Students; All Covariates	-5.676	3.722	-1.525	0.129
Pre Students; Program Only	3.990	4.687	0.851	0.396
Pre Students; Program & School	2.997	5.014	0.598	0.551
Pre Students; Key Covariates	3.899	4.992	0.781	0.436
Pre Students; All Covariates	2.500	5.112	0.489	0.626
Pre; Program & Both Pretests	5.848	4.485	1.304	0.194
Pre; Program, School, Both Pre	5.086	4.777	1.065	0.289
Pre; Key Covariates & Both Pre	5.838	4.755	1.228	0.222
Pre; All Covariates & Both Pre	4.482	4.894	0.916	0.361
Pre; Program & Pretest Total	5.138	4.557	1.128	0.261
Pre; Program, School Pre Total	4.740	4.881	0.971	0.333
Pre; Key Covariates & Pre Total	5.390	4.876	1.105	0.271
Pre; All Covariates & Pre Total	3.651	4.982	0.733	0.465
1-3 Students; Program Only	-6.458	4.134	-1.562	0.120
1-3 Students; Program & School	-7.627	4.077	-1.870	0.063
1-3 Students; Key Covariates	-6.779	4.099	-1.654	0.100
1-3 Students; All Covariates	-7.855	4.084	-1.923	0.056

The K-1 mathematics analyses on students with pretest scores showed an advantage for immersion strategy students, but the effect is not even close to statistical significance. Separate analyses revealed that the reversal of the direction of the program difference is largely attributable to the elimination of school 51, whose students never had pretest

ERIC

scores. Although program effect changes from school to school imply a program-by-school interaction, formal tests for such interactions were not statistically significant. Thus the overall conclusion is that there are no significant differences between the two programs in mathematics at spring of first grade.

In addition, the school differences are attributable to school 51, which has higher mathematics achievement than the other schools. Covariate-adjusted estimates of the school effects revealed that the other three schools all had approximately equal mathematics achievement levels.

The findings are essentially unchanged by the inclusion of pretest scores as covariates.

What were the conclusions from the K-1 analyses for the language subtest?

As for the mathematics subtest, the K-1 analyses of the language subtest (Table 87) showed no statistically significant differences between the immersion strategy and early-exit students. However, the sign and magnitude of the program difference is consistent across all the analyses, showing a small immersion strategy advantage. School effects were statistically significant, with school 51 showing higher achievement, schools 12 and 14 lower achievement, and school 21 about average. The reduction of the sample to students with pretest scores and the inclusion of pretest scores as covariates slightly strengthened the IS advantage, but the program difference still did not reach statistical significance.

Table 87

Two-Program Schools K-1 Analysis:
Estimated Program Differences in Spring 1st for Language

Students; Analysis	Parameter	Std Err	t-Stat	p-Value
All Students; Program Only	5.185	4.726	1.097	0.274
All Students; Program & School	4.019	4.651	0.864	0.388
All Students; Key Covariates	5.004	4.638	1.079	0.282
All Students; All Covariates	5.222	4.691	1.113	0.267
Pre Students; Program Only	6.675	6.030	1.107	0.270
Pre Students; Program & School	7.596	6.456	1.177	0.241
Pre Students; Key Covariates	7.984	6.464	1.235	0.219
Pre Students; All Covariates	8.263	6.720	1.230	0.221
Pre; Program & Both Pretests	7.648	5.921	1.292	0.198
Pre; Program, School, Both Pre	9.507	6.330	1.502	0.135
Pre; Key Covariates & Both Pre	9.757	6.350	1.537	0.127
Pre; All Covariates & Both Pre	9.298	6.597	1.409	0.161
Pre; Program & Pretest Total	7.987	5.906	1.352	0.178
Pre; Program, School Pre Total	9.651	6.323	1.526	0.129
Pre; Key Covariates & Pre Total	9.920	6.342	1.564	0.120
Pre; All Covariates & Pre Total	9.671	6.577	1.470	0.144
1-3 Students; Program Only	7.544	5.244	1.438	0.152
1-3 Students; Program & School	7.168	5.122	1.400	0.163
1-3 Students; Key Covariates	8.432	5.156	1.635	0.104
1-3 Students; All Covariates	8.589	5.210	1.649	0.101

What were the conclusions from the K-1 analyses for the English reading subtest?

The K-1 analyses of the English reading subtest are summarized in Table 88. The analyses of all students showed a statistically significant advantage for early-exit students. The early-exit advantage was more pronounced among the students with pretests (and therefore excluding students from school 51). The inclusion of pretest scores as covariates slightly weakened the EE advantage, but the program difference remained clearly significant, at the .01 level. Among the students in the 1-3 analyses, however, the program difference was slightly smaller and no longer achieved statistical significance. School effects were generally statistically significant, with school 51 showing higher achievement and school 14 lower achievement.

Table 88

Two-Program Schools K-1 Analysis:
Estimated Program Differences in Spring 1st for Reading

Students; Analysis	Parameter	Std Err	t-Stat	p-Value
All Students; Program Only	-7.147	3.859	-1.852	0.065
All Students; Program & School	-8.977	3.792	-2.367	0.019
All Students; Key Covariates	-8.177	3.773	-2.167	0.031
All Students; All Covariates	-8.433	3.793	-2.223	0.027
Pre Students; Program Only	-13.086	4.901	-2.670	0.008
Pre Students; Program & School	-14.932	5.218	-2.861	0.005
Pre Students; Key Covariates	-14.619	5.232	-2.794	0.006
Pre Students; All Covariates	-16.040	5.347	-3.000	0.003
Pre; Program & Both Pretests	-12.677	4.880	-2.598	0.010
Pre; Program, School, Both Pre	-13.924	5.205	-2.675	0.008
Pre; Key Covariates & Both Pre	-13.707	5.226	-2.623	0.010
Pre; All Covariates & Both Pre	-15.552	5.326	-2.920	0.004
Pre; Program & Pretest Total	-12.360	4.871	-2.537	0.012
Pre; Program, School Pre Total	-13.817	5.197	-2.659	0.009
Pre; Key Covariates & Pre Total	-13.586	5.216	-2.605	0.010
Pre; All Covariates & Pre Total	-15.259	5.310	-2.874	0.005
1-3 Students; Program Only	-5.165	4.317	-1.196	0.233
1-3 Students; Program & School	-6.585	4.243	-1.552	0.122
1-3 Students; Key Covariates	-4.784	4.233	-1.130	0.260
1-3 Students; All Covariates	-5.459	4.234	-1.289	0.199

Results of the 1-3 Analyses

The kindergarten through first grade analyses were completed to determine whether immersion strategy and early-exit students had comparable skills in mathematics, English language and English reading in spring of first grade. The analyses that follow were carried out to compare the growth in these skills from first grade through third grade between immersion strategy and early-exit students. The analyses are directed towards determining if the growth rate is comparable between students in the two programs.

What were the results of the 1-3 analysis for the mathematics subtest?

For the two-program schools analyses, first through third grade information was used to estimate individual growth curves. For each student with test scores at two or more time points, an individual growth curve was fit to the subtest scores for each subtest (within-student growth curve). An overall coefficient of curvature was estimated and then an intercept (status at spring of first grade) and slope (growth rate from spring of first grade through spring of third grade) was estimated for each individual. These individual intercepts and slopes were then predicted using background variables, a set of variables (sometimes termed "dummy variables") for indicating the school, and a single binary variable for the difference between the immersion strategy and early-exit programs. That is, each student is allowed to have their own growth curve (within-student analysis). This information is used to fit a model of an overall growth curve between groups of students across programs (between-student growth curves). The school and program variables are included as predictors of both the initial status and the growth rate.

Table 89 presents the estimated coefficients, their estimated standard errors, and an approximate t-statistic and associated p-value for the mathematics subtest. This table will be described in detail. The discussion of other tables of HLM results, including Table 90 (for the language subtest) and Table 91 (for the reading subtest), will be limited to pointing out highlights.

Table 89

Two-Program Schools 1-3 Analysis for Mathematics

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	247.184	5.131	48.174	0.000
Spring 1st	PGMIS	-5.710	3.637	-1.570	0.118
Spring 1st	SCHOOL12	-3.245	2.715	-1.195	0.233
Spring 1st	SCHOOL14	-3.436	3.175	-1.082	0.281
Spring 1st	SCHOOL51	9.790	3.566	2.745	0.007
Spring 1st	AVABS13	-0.114	0.268	-0.425	0.671
Spring 1st	FEMALE	-7.264	3.387	-2.145	0.033
Spring 1st	PRESCHY	8.110	3.231	2.510	0.013
Spring 1st	EDAVG	1.389	0.511	2.719	0.007
Growth Rate 1-3	BASE	35.822	3.259	10.993	0.000
Growth Rate 1-3	PGMIS	1.428	2.765	0.517	0.606
Growth Rate 1-3	SCHOOL12	-0.274	2.213	-0.124	0.902
Growth Rate 1-3	SCHOOL14	3.472	2.526	1.375	0.171
Growth Rate 1-3	SCHOOL51	6.221	2.680	2.321	0.021
Growth Rate 1-3	AVABS13	-0.534	0.226	-2.365	0.019
Growth Rate 1-3	FEMALE	7.178	2.540	2.826	0.005
Curvature 1-3	BASE	-6.959	1.713	-4.062	0.000

The first column of Table 89, labeled "Predicted", indicates what aspect of the achievement test Expanded Scale Scores is being predicted. The "Spring 1st" lines show the predictors of the estimated ESS at time zero (spring of first grade), also termed initial status or intercept. Recall that initial status is estimated separately for each student. An increase in the initial status shifts the entire growth curve upward without changing its slope or curvature, while a decrease in the initial status shifts the entire growth curve downward, again without changing its slope or curvature.

The "Growth Rate 1-3" lines show the predictors of the average annual increase in test score from spring of first grade to spring of third grade. The growth rate or slope is also estimated separately for each student. An increase in the growth rate tilts the growth curve to a

ERIC

steeper angle, while a decrease in the growth rate tilts the growth curve to a flatter angle, without changing the amount of curvature.

The "Curvature 1-3" line indicates the curvature of the growth from spring of first grade to spring of third grade. The coefficient of curvature is the same for all students, so the table shows only one line for curvature. As noted in Chapter II, there are not enough time points to estimate a separate coefficient of curvature for each individual. The negative coefficient for curvature indicates a curve that tends to flatten, which matches the shape of the curve for the national norms.

The "Predictor" column gives the name of the variable used as a predictor. In addition to some of the background variables described in Chapter II, this column includes BASE, PGMIS, SCHOOL12, SCHOOL14, and SCHOOL51. The predictor labeled BASE is a constant term in the prediction equation. The term intercept will not be used to refer to this constant to avoid confusion with the intercept (initial status) of the individual growth curves. PGMIS is a binary variable that is 1 for students in the immersion strategy program and 0 for students in the early-exit program.

The SCHOOL variables are indicators of the child's school. For example, a child in school 12 would have the value 1 for SCHOOL12 and the value 0 for the other two SCHOOL variables. Children in school 21 were given the value -1 for all three school variables, making school 21 the "omitted school" in this parameterization. The parameters for the SCHOOL variables represent the difference between that school and the average of all four schools. The difference between the omitted school, school 21, and the average of all four schools can be calculated as the negative of the sum of the coefficients for the three SCHOOL variables. Any school could have been selected as the omitted school; the choice does not affect the statistical results.

The interpretation of the parameter value for AVABS13 in predicting the initial status of ESS scores (Spring 1st) is that for every day absent the child's predicted score is lowered by .114 ESS points. Thus a child absent an average of 10 days a year is expected to score about 1.14 points lower than that same child would have scored if not absent at all. Since the predicted initial status is adjusted for the effect of absences, the estimated initial status for such a child is 1.14 points greater than the test scores (which reflect the 10 days of absence every year) would indicate.

The estimated standard error of the parameter for AVABS13 is 0.268. The coefficient for AVABS13 (that is, the parameter value associated with AVABS13), -0.114, divided by its standard error, 0.268, is an approximate t-statistic (-0.425) which is not statistically significant (p = .671). The variable is included even though not statistically significant in order to facilitate comparison with other models and to ensure that the results are not distorted by differential absences across schools or programs. The "p-Value" given is for a two-tailed test.

The predicted initial status for a given child can be calculated from the parameter values in the table and the value of the predictors for that child. For example, if the child attended preschool, PRESCHY would have the value 1. If the child's parents were both high school graduates, EDAVG would have the value 12. The predicted initial status for a boy with these values who was in the early-exit program in school 12 and never absent would be calculated as (247.184) + (-3.245 x 1) + (8.110 x 1) + (1.389 x 12) = 268.717. A boy with the same values but in the immersion strategy program in school 12 would have a predicted initial status 5.710 points lower, or 263.007.

The calculation is somewhat more complicated for students in the omitted school, school 21. These students have a value of -1 for all the SCHOOL variables. Thus the predicted initial status for a boy in the early-exit program in school 21 who was never absent, attended preschool, and with EDAVG of 12 would be calculated as (247.184) + (-3.245 x -1) + (-3.436 x -1) + (9.790 x -1) + (8.110 x 1) + (1.389 x 12) = 268.853. Similar calculations can be made for other students.

Thus the parameters for the three SCHOOL variables indicate the increase (or decrease) in the predicted value for a school relative to the average of the four schools. The corresponding figure for school 21 is obtained by negating the sum of the coefficients for the three SCHOOL variables. This means the corresponding initial status coefficient for school 21 is estimated to be -(-3.245 + -3.436 + 9.790) = -3.109. That is, school 21 is estimated to have an initial status 3.109 points lower than the average of all four schools.

The parameter value for PGMIS indicates the increase (or decrease) in the predicted value for a student in an immersion strategy program. The value -5.710 for the parameter PGMIS for predicting initial status means the initial status of immersion strategy students is predicted to be about 5.7 ESS points below that of early-exit students in the same school. This estimate is the same for (and based on the information from) all four two-program schools. It is not statistically significant (p = .118), so one cannot reject the null hypothesis that there is no program difference in mathematics achievement in spring of first grade based on the 1-3 analysis.

The t-statistic and associated p-value for each SCHOOL variable tests the hypothesis that the school indicated has a parameter value of zero. This is equivalent to the hypothesis that the early-exit program in this school has the same initial status as the average of the early-exit programs in the four schools, and that the immersion strategy program in this school has the same initial status as the average of the immersion strategy programs in the four schools. Schools 12 (p = .233) and 14 (p = .281) have approximately the same initial status as the average of the four schools. The omitted school 21 can be tested using a multivariate hypothesis test, which indicates that it is not significantly different from the average, either (t = -0.662, p = .508). Since the difference between immersion strategy and early-exit is modeled to be the same in all IS/EE schools, this implies that the early-exit programs in 12, 14, and 21 have approximately the same initial status, as do their immersion strategy programs. However, school 51 (p = .007) has significantly higher initial status than the average of the four schools.

An overall test that the four schools differ significantly on initial status is obtained from a multivariate hypothesis test (chi-square with 3 d.f. = 8.9172, p = .030). Note that the overall test may be significant when no individual school differs significantly from the average of the four schools and the overall test may be nonsignificant even though an individual school does appear to differ significantly from the average. A conservative approach is to consider the individual school tests only when the overall test of school differences is statistically significant.

The table does not indicate whether, for example, the early-exit program in school 12 (parameter value for SCHOOL12 -3.245, standard error 2.715) differs from the immersion strategy program in school 51 (parameter value for SCHOOL51 9.790, standard error 3.566, plus the parameter value for PGMIS -5.710, standard error 3.637). The estimated difference can be calculated as (9.790 + -5.710) - (-3.245) = 7.325 but the standard error of that difference cannot be obtained from the table because of correlations among the parameter estimates. By referring to the standard error of each parameter, however, it is possible to get a sense of whether the estimated difference is large or small. In this case the estimated difference of 7.325 is about twice the standard error of the individual parameters, so the difference is moderate. A difference three or four times the standard errors is large, while a difference smaller than the standard errors is small. These guidelines should not be relied upon, however, for drawing conclusions.

The prediction equation for the growth rate is interpreted similarly to the prediction equation for the initial status. Notice that the only background variables found to be predictive of mathematics growth rate are AVABS13, the average number of days the child was absent from school during the study, and FEMALE, whether the child is a girl.

The coefficient of AVABS13 in predicting the growth rate, -0.534, indicates that each additional day of absence (averaged across the study) reduces the annual growth rate by .534 points. Interpreted in connection with the coefficient of -0.114 for AVABS13 in predicting the initial status, this means that absences not only shift the growth curve downward by about .114 points per day of absence (averaged) but also tend to make the growth curve less steep by about .534 points per day of absence (averaged). Because the number of absences is a predictor of the growth rate as well as the overall level of the curve, absences can be thought of as having a cumulative effect. For example, a child absent exactly 10 days every year would have a growth curve shifted downward by 1.14 points and tilted so that the growth was reduced by 5.34 points every year. At time 2 (spring of third grade), a child absent 10 days every year would have a predicted value that is (10 x .114) + (10 x .534 x 2) = 10.68 below the value that would be predicted with no absences.

The growth rate for girls is estimated to be 7.178 points higher than for boys, whereas the initial status for girls is estimated to be 7.264 points lower than for boys. Taken together, these estimates indicate that although girls lag behind boys in mathematics at spring of first grade, boys and girls are about equal by spring of second grade and girls are ahead by the end of third grade.

The mathematics growth rate for the immersion strategy program is essentially the same as for the early-exit program (difference = 1.428, p = .606). The growth rates for the four schools show marginally significant variation (chi-square with 3 d.f. = 7.778, p = .051). The growth rate for the omitted school 21 is lowest, 9.419 points below the average of the four schools: -(-0.274 + 3.472 + 6.221) = -9.419. This is significantly different from zero (p = .044), indicating that school 21 has a lower than average growth rate. School 51 has the highest growth rate, 6.221 above average (p = .021). Schools 12 and 14 do not differ significantly from the average.

ERIC

Only the BASE is shown as a predictor for curvature. Although the possibility that curvature varied according to background variables, school, and program was explored, no statistically significant variation was found. In particular, there were no significant program differences in the amount of curvature.

Summary

For the mathematics subtest there were no statistically significant differences between the immersion strategy and early-exit programs in the growth curves. Both curves indicate that the rate of growth is decreasing over time (i.e., flattening out). There were differences among the four schools, with school 51 showing the highest initial status and the highest growth rate among the four schools.

What were the results of the 1-3 analysis for the language subtest?

Table 90 presents the two-program schools analysis for the language subtest. The interpretation of this table is similar to the interpretation of Table 89, the two-program schools results for the mathematics subtest. The predictors of initial status for the language subtest differ considerably from the predictors for the mathematics subtest. In addition, only FEMALE among the background variables was found to be a significant predictor of the growth rate for language.

Table 90

Two-Program Schools 1-3 Analysis for Language

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	218.783	7.004	31.235	0.000
Spring 1st	PGMIS	11.037	5.245	2.104	0.037
Spring 1st	SCHOOL12	-6.363	3.840	-1.657	0.099
Spring 1st	SCHOOL14	-10.553	4.608	-2.290	0.023
Spring 1st	SCHOOL51	8.539	5.171	1.651	0.100
Spring 1st	AVABS13	0.166	0.331	0.503	0.616
Spring 1st	ANYEPTOP	13.837	6.732	2.056	0.041
Spring 1st	EDAVG	1.930	0.790	2.443	0.015
Spring 1st	BOOKSHM	2.461	1.577	1.560	0.120
Growth Rate 1-3	BASE	53.044	5.112	10.376	0.000
Growth Rate 1-3	PGMIS	-4.304	4.857	-0.886	0.377
Growth Rate 1-3	SCHOOL12	5.523	4.020	1.374	0.171
Growth Rate 1-3	SCHOOL14	2.792	4.571	0.611	0.542
Growth Rate 1-3	SCHOOL51	15.359	4.747	3.235	0.001
Growth Rate 1-3	FEMALE	10.881	3.841	2.833	0.005
Curvature 1-3	BASE	-10.253	3.316	-3.092	0.002

The immersion strategy programs have higher initial status than the early-exit programs (difference = 11.037, p = .037), but the growth rates do not differ significantly (difference = -4.304, p = .377). Because the growth rates are not significantly different, one cannot reject the hypothesis that the higher achievement among IS students at spring of first grade is maintained through third grade. However, notice that the estimated growth rate for the immersion strategy programs is lower than the estimated growth rate for the early-exit programs (albeit not significantly lower). This implies that the early-exit students may be narrowing the gap. In fact, the estimated difference between the immersion strategy and the early-exit students at the end of third grade (time 2) is only 11.037 + 2 x (-4.304) = 2.429. A separate test shows that this difference is not statistically significant (p = .768). Thus one cannot reject the hypothesis that the immersion strategy and early-exit programs have the same level of language achievement at third grade.

School differences exist for both initial status (p = .030) and for growth rate (p = .010). As for mathematics, school 51 exhibits a significantly high growth rate relative to the other schools (difference 15.359, p = .001), but the somewhat higher initial status is not significant (difference from the average of 8.539, p = .100). The higher initial status for the omitted school 21 is also not significant (difference 8.377, p = .222), but again, as was found for mathematics, school 21 has a significantly lower growth rate (difference -23.674, p = .006).

Usually any background variable included as a predictor of the growth rate is included as a predictor of initial status even if it does not achieve statistical significance. For the language subtest, however, FEMALE was not included as a predictor of initial status even though it was a highly significant predictor of growth rate, with girls' growth rates estimated to be nearly 11 points higher than boys'. FEMALE was omitted from the Spring 1st prediction equation because the estimated effect was minuscule, indicating that boys and girls had nearly identical estimated initial status.

It is noteworthy that the sign of AVABS13 is positive as a predictor of initial status for the language subtest. This implies that absences actually improve the student's score on the language subtest. However, the value is not significantly different from zero (p = .616) and is presumably acting as a surrogate for other variables. The variable was included despite "having the wrong sign" for ease of comparisons with other models.

Summary

For the language subtest there is evidence that the immersion strategy students have higher achievement at spring of first grade than early-exit students. However, there is no significant difference between the immersion strategy and early-exit students by the end of third grade. That the base of the slope is significant suggests that students in both programs are growing. Although the growth rates are not statistically significantly different, numerically there is enough difference to allow for early-exit students to catch up to immersion strategy students. Any relative advantage of the IS program for the language subtest appears to be fairly small and unsustained.

What were the results of the 1-3 analysis for the English reading subtest?

Table 91 gives the two-program schools analysis for the English reading subtest. The predictors of initial status for the reading subtest are the same as for the language subtest. The coefficients for the background variables ANYEPTOP, EDAVG, and BOOKSHM have the same sign in the language and reading analyses and the magnitudes are similar for ANYEPTOP and BOOKSHM. For EDAVG, the estimated coefficient is considerably smaller for the reading subtest and does not achieve statistical significance ($p = .368$). Notice that although BOOKSHM has similar magnitude in the language and reading analyses, the small difference in estimated standard error leaves the coefficient nonsignificant for language ($p = .120$) but significant for reading ($p = .045$).

Table 91

Two-Program Schools 1-3 Analysis for English Reading

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	221.766	5.697	38.926	0.000
Spring 1st	PGMIS	1.406	4.322	0.325	0.745
Spring 1st	SCHOOL12	0.346	3.126	0.111	0.912
Spring 1st	SCHOOL14	-7.645	3.751	-2.038	0.043
Spring 1st	SCHOOL51	7.138	4.207	1.697	0.091
Spring 1st	AVABS13	-0.197	0.268	-0.736	0.463
Spring 1st	ANYEPTOP	14.095	5.450	2.586	0.010
Spring 1st	EDAVG	0.575	0.638	0.902	0.368
Spring 1st	BOOKSHM	2.581	1.278	2.020	0.045
Growth Rate 1-3	BASE	52.896	3.828	13.817	0.000
Growth Rate 1-3	PGMIS	-1.689	4.034	-0.419	0.676
Growth Rate 1-3	SCHOOL12	-0.459	3.195	-0.144	0.886
Growth Rate 1-3	SCHOOL14	5.098	3.638	1.401	0.163
Growth Rate 1-3	SCHOOL51	9.550	3.773	2.531	0.012
Curvature 1-3	BASE	-4.394	4.800	-0.915	0.361
Curvature 1-3	PGMIS	-17.543	5.687	-3.085	0.002

Unlike the results for language, FEMALE was not found to be a statistically significant predictor of the English reading growth rate. Also unlike language, neither the programs (p = .745) nor the schools as a group (p = .125) were found to differ in initial status. For the growth rate, however, the absence of program differences (p = .676) and the presence of school differences (p = .032) matches the findings for language. Once again, school 51 was found to have a relatively high growth rate (difference 9.550, p = .012) and omitted school 21 a relatively low growth rate (difference -14.189, p = .037).

By far the most striking difference between the results for the reading subtest and the results for the mathematics and language subtests, however, is the finding that the amount of curvature differs according to program. The curvature was not found to vary according to school or background variables. In addition, the difference in curvature between the two programs did not allow the curvature to vary according to other variables. The difference between the programs is that the curvature is more negative for IS than it is for EE. This means that students in both programs tend to have growth rates that tend to flatten, but the IS students tend to have higher growth at first and lower growth at the end. Interpreted in connection with the small and nonsignificant differences between the programs in initial status and growth rate, this means that the two programs have essentially the same average achievement and the same average growth rate from first to third grades, but that immersion strategy students have higher reading achievement at second grade and therefore lower reading achievement at first and third grades.

To see how this follows from the estimates of program differences for the initial status, growth rate, and curvature, consider the estimated difference between IS and EE in spring of each grade. At spring of first grade (initial status), IS is estimated to be 4.442 points behind EE, calculated as 1.406 + (1/3) x (-17.543). At spring of second grade, IS is estimated to be 11.412 points ahead of EE, calculated as 1.406 + 1 x (-1.689) + (-2/3) x (-17.543). And at spring of third grade, IS is estimated to be 7.820 points behind EE, calculated as 1.406 + 2 x (-1.689) + (1/3) x (-17.543).

The estimated program differences at first grade (-4.442) and at third grade (-7.820) are not significant ($p = .318$ and $p = .293$, respectively). However, the program difference at second grade (11.412) is significant ($p = .024$).

Summary

The English reading subtest results show an advantage for the immersion strategy programs at second grade, but no significant difference between the programs at first or third grade. At third grade, the IS students have lower growth rates than the EE students as a consequence of the estimated curvature difference. In addition, the IS students have slightly lower achievement at spring of third grade but the difference is not statistically significant. These results are similar to the finding for the language subtest in the sense that IS students have a small advantage at one grade that appears not to be sustained. A difference between the two subtests is that for language the IS advantage appears for the test in spring of first grade, whereas for reading the IS advantage appears for the test in spring of second grade.

Including the Pretest as a Predictor in the 1-3 Analyses

How were the 1-3 analyses for mathematics affected by controlling for pretest?

Just as the K-1 analyses were repeated to see the effect of controlling for the pretest, so were the 1-3 analyses. Again, before evaluating the effect of including the pretest as a covariate it is necessary to evaluate the effect of reducing the students analyzed to those who have a pretest. This means leaving out all school 51 students. Table 92 summarizes the resulting HLM model for the mathematics subtest. This table should be compared with Table 89.

ERIC

Table 92

Two-Program Schools: 1-3 Analysis for Mathematics
Students with Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	236.903	7.425	31.908	0.000
Spring 1st	PGMIS	0.619	5.996	0.103	0.918
Spring 1st	SCHOOL12	-3.043	3.716	-0.819	0.415
Spring 1st	SCHOOL14	1.565	3.875	0.404	0.687
Spring 1st	AVABS13	0.103	0.323	0.317	0.752
Spring 1st	FEMALE	-10.412	4.867	-2.139	0.035
Spring 1st	PRESCHY	9.230	4.505	2.049	0.043
Spring 1st	EDAVG	1.855	0.736	2.521	0.013
Growth Rate 1-3	BASE	32.330	5.984	5.403	0.000
Growth Rate 1-3	PGMIS	-3.046	5.675	-0.537	0.593
Growth Rate 1-3	SCHOOL12	4.776	3.323	1.437	0.154
Growth Rate 1-3	SCHOOL14	2.195	3.492	0.628	0.531
Growth Rate 1-3	AVABS13	-0.639	0.319	-2.003	0.048
Growth Rate 1-3	FEMALE	9.187	4.282	2.145	0.034
Curvature 1-3	BASE	-13.940	2.165	-6.440	0.000

* Curvature 1-3: PGMIS omitted; not statistically significant.

Among all students, the EE students showed a small and nonsignificant advantage in initial status (spring of first grade), but the IS students showed a very small and nonsignificant advantage in growth rate. For the students with a pretest, the three remaining schools show a program difference of less than one point for initial status (difference = 0.619, p = .918) and a small early-exit advantage in growth rate (difference = -3.046) that is not near statistical significance (p = .593). Thus although the sign of the program differences are changed, they all remain nonsignificant. The conclusion that mathematics achievement is essentially the same for the two programs is sustained.

The school differences for initial status and for the growth rate among the three schools are both nonsignificant (p = .708 and p = .242, respectively). The coefficient of AVABS13 as a predictor of initial status is still nonsignificant but has changed sign. The other predictors of initial status and growth rate are similar in sign and magnitude. The coefficient of curvature is somewhat more negative, however, indicating a somewhat greater amount of curvature for these students.

Table 93 presents the results of including the mathematics pretest (ANALPREM) as a predictor of initial status. The language pretest was found not to predict initial status and neither pretest was found to predict growth rate. Although the IS advantage at initial status is increased to 3.853 points in this analysis, the difference is not close to statistical significance (p = .510). The EE advantage in growth rate is little changed and still not statistically significant (difference = -3.434, p = .544). The other parameters show remarkably little change as a result of including the mathematics pretest as a predictor.

Table 93

Two-Program Schools 1-3 Analysis for Mathematics
Mathematics Pretest as Covariate

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	150.097	21.448	6.998	0.000
Spring 1st	PGMIS	3.853	5.825	0.662	0.510
Spring 1st	SCHOOL12	-3.545	3.563	-0.995	0.322
Spring 1st	SCHOOL14	0.171	3.742	0.046	0.964
Spring 1st	AVABS13	0.054	0.312	0.174	0.862
Spring 1st	FEMALE	-10.905	4.683	-2.329	0.022
Spring 1st	PRESCHY	9.218	4.188	2.201	0.030
Spring 1st	EDAVG	1.231	0.699	1.761	0.081
Spring 1st	ANALPREM	0.553	0.129	4.284	0.000
Growth Rate 1-3	BASE	32.810	5.952	5.512	0.000
Growth Rate 1-3	PGMIS	-3.434	5.646	-0.608	0.544
Growth Rate 1-3	SCHOOL12	4.632	3.307	1.401	0.164
Growth Rate 1-3	SCHOOL14	2.120	3.475	0.610	0.543
Growth Rate 1-3	AVABS13	-0.606	0.318	-1.909	0.059
Growth Rate 1-3	FEMALE	8.856	4.260	2.079	0.040
Curvature 1-3	BASE	-13.709	2.152	-6.369	0.000

* Curvature 1-3: PGMIS omitted; not statistically significant.

Summary

For the mathematics subtest the finding of no statistically significant differences between the immersion strategy and early-exit programs was not altered by the analyses of students with pretest scores. Furthermore, the omission of students in school 51 (none of whom had pretest scores) eliminated the school differences in initial status and growth rate.

How were the 1-3 analyses for language affected by controlling for pretest?

The HLM model for the language subtest for 1-3 students with pretests is presented in Table 94. For all students, the 1-3 analysis in Table 90 showed a statistically significant immersion strategy advantage in initial status (difference = 11.037, p = .037). For the students with pretest scores, the immersion strategy advantage is increased to 17.645 (p = .020). For all students, immersion strategy students showed a lower growth rate (difference = -4.304) which, although nonsignificant (p = .377), had the effect of eroding the advantage in spring of first grade. The estimated program difference at third grade was not significant (difference = 2.429, p = .768). For the students with pretest scores, the program difference in initial status is larger and the difference in growth rate is smaller (difference = -2.173, p = .796). However, the program difference at spring of third grade is still not significant (difference = 13.299, p = .372).

Table 94

Two-Program Schools 1-3 Analysis for Language
Students With Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	227.888	9.133	24.952	0.000
Spring 1st	PGMIS	17.645	7.484	2.358	0.020
Spring 1st	SCHOOL12	-1.263	4.428	-0.285	0.776
Spring 1st	SCHOOL14	-4.499	4.876	-0.923	0.358
Spring 1st	AVABS13	-0.014	0.363	-0.039	0.969
Spring 1st	ANYEPTOP	5.877	10.654	0.552	0.582
Spring 1st	EDAVG	1.251	1.114	1.123	0.264
Spring 1st	BOOKSHM	0.342	2.095	0.163	0.871
Growth Rate 1-3	BASE	36.240	8.216	4.411	0.000
Growth Rate 1-3	PGMIS	-2.173	8.394	-0.259	0.796
Growth Rate 1-3	SCHOOL12	15.417	5.018	3.072	0.003
Growth Rate 1-3	SCHOOL14	4.461	5.245	0.851	0.397
Growth Rate 1-3	FEMALE	13.192	5.451	2.420	0.017
Curvature 1-3	BASE	-25.233	3.941	-6.402	0.000

* Curvature 1-3: PGMIS omitted; not statistically significant.

The school differences for initial status among the three schools are nonsignificant (p = .572), but the school differences for growth rate are significant (p = .005). The growth rate for the omitted school, school 21, is significantly lower than the average for the three schools (difference = -19.878, p = .005) and the growth rate for school 12 is significantly higher (difference = 15.417, p = .003). The coefficient of AVABS13 as a predictor of initial status is nearly zero, and the coefficient of ANYEPTOP is reduced in magnitude and has much higher standard error compared with the 1-3 analysis of language using all students. The other predictors of initial status and growth rate are similar in sign and magnitude. The estimated growth rates for schools 14 and 21 are lower, and the amount of negative curvature is greater among these students with pretest scores in schools 12, 14 and 21 than among all students in the four two-program schools. The estimated standard errors of the growth rates are substantially increased, however, indicating that the differences between the analyses are probably not significant.

Table 95 presents the results of including the sum of the mathematics and language subtests (ANALTOBE) as a predictor of initial status. The sum of the pretests was found to be a slightly better predictor of initial status than the language pretest and a much better predictor than the mathematics pretest. Neither one of the pretests alone nor their sum was a significant predictor of the growth rate. Although the IS advantage at initial status is increased again, to 20.317 points (p = .008), and the program difference in growth rate is still small (difference = -2.876, p = .730), the estimated program difference at third grade once again is not statistically significant (difference = 14.565, p = .319). The other parameters are little affected by the inclusion of pretest as a predictor.

Table 95

Two-Program Schools 1-3 Analysis for Language
Pretest Total as Covariate

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	160.529	26.438	6.072	0.000
Spring 1st	PGMIS	20.317	7.488	2.713	0.008
Spring 1st	SCHOOL12	-2.115	4.402	-0.480	0.632
Spring 1st	SCHOOL14	-5.147	4.843	-1.063	0.290
Spring 1st	AVABS13	-0.018	0.357	-0.049	0.961
Spring 1st	ANYEPTOP	3.700	10.487	0.353	0.725
Spring 1st	EDAVG	0.890	1.103	0.808	0.421
Spring 1st	BOOKSHM	0.535	2.058	0.260	0.795
Spring 1st	ANALTOBE	0.206	0.076	2.699	0.008
Growth Rate 1-3	BASE	37.611	8.123	4.630	0.000
Growth Rate 1-3	PGMIS	-2.876	8.306	-0.346	0.730
Growth Rate 1-3	SCHOOL12	15.101	4.973	3.037	0.003
Growth Rate 1-3	SCHOOL14	4.524	5.192	0.871	0.386
Growth Rate 1-3	FEMALE	11.856	5.351	2.216	0.029
Curvature 1-3	BASE	-25.109	3.924	-6.398	0.000

* Curvature 1-3: PGMIS omitted; not statistically significant.

Summary

The finding of higher immersion strategy achievement on the language subtest at spring of first grade was confirmed and strengthened by the analysis of students with pretest scores. Even though the estimated program difference at third grade was larger among the students with pretest scores, it was not close to statistical significance. The three schools with pretest information exhibited no statistically significant differences in initial status but did have statistically significant growth rate differences.

How were the 1-3 analyses for reading affected by controlling for pretest?

Table 96 gives the reading model for the students with pretests. Like the model for all students, this model shows no statistically significant program difference in initial status ($p = .500$) or in growth rate ($p = .614$), but it does show a significant program difference in curvature ($p = .002$). Based on the estimated program differences in initial status, growth rate, and curvature, the estimated program difference at spring of first grade is -11.264 ($p = .093$). The estimated program difference at spring of second grade is 12.110 ($p = .106$) and the estimated program difference at spring of third grade is -4.274 ($p = .746$). These differences are similar to those found in the analysis of all students, with the second grade advantage for IS students no longer achieving statistical significance.

Table 96

Two-Program Schools 1-3 Analysis for Reading
Students With Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	229.348	8.226	27.881	0.000
Spring 1st	PGMIS	-4.638	6.849	-0.677	0.500
Spring 1st	SCHOOL12	3.981	3.934	1.012	0.314
Spring 1st	SCHOOL14	-1.975	4.326	-0.456	0.649
Spring 1st	AVABS13	-0.249	0.327	-0.763	0.447
Spring 1st	ANYEPTOP	10.153	9.579	1.060	0.292
Spring 1st	EDAVG	0.589	1.000	0.589	0.557
Spring 1st	BOOKSHM	3.038	1.885	1.612	0.110
Growth Rate 1-3	BASE	39.321	5.924	6.638	0.000
Growth Rate 1-3	PGMIS	3.495	6.915	0.505	0.614
Growth Rate 1-3	SCHOOL12	5.770	3.762	1.534	0.128
Growth Rate 1-3	SCHOOL14	3.352	3.944	0.850	0.397
Curvature 1-3	BASE	-11.094	6.650	-1.668	0.098
Curvature 1-3	PGMIS	-19.879	7.449	-2.669	0.009

The school differences for initial status and growth rate are nonsignificant ($p = .586$ and $p = .201$, respectively). The coefficients for the covariates generally exhibit the same pattern as for the analysis of the language subtest.

The sum of the mathematics and reading subtests (ANALTOBE) was included as a predictor of initial status in reading (Table 97). As for the language subtest, the sum of the pretests was found to be a slightly better predictor of initial status than the language pretest and a much better predictor than the mathematics pretest. Neither pretest alone nor their sum was a significant predictor of the growth rate.

Table 97

Two-Program Schools 1-3 Analysis for Reading
Pretest Total as Covariate

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	184.834	23.714	7.794	0.000
Spring 1st	PGMIS	-2.879	6.889	-0.418	0.677
Spring 1st	SCHOOL12	3.386	3.934	0.861	0.392
Spring 1st	SCHOOL14	-2.428	4.323	-0.562	0.576
Spring 1st	AVABS13	-0.257	0.323	-0.797	0.428
Spring 1st	ANYEPTOP	8.775	9.487	0.925	0.357
Spring 1st	EDAVG	0.382	0.996	0.383	0.702
Spring 1st	BOOKSHM	3.105	1.862	1.667	0.099
Spring 1st	ANALTOBE	0.136	0.069	1.986	0.050
Growth Rate 1-3	BASE	39.446	5.883	6.705	0.000
Growth Rate 1-3	PGMIS	3.419	6.867	0.498	0.620
Growth Rate 1-3	SCHOOL12	5.657	3.749	1.509	0.134
Growth Rate 1-3	SCHOOL14	3.363	3.926	0.857	0.394
Curvature 1-3	BASE	-11.047	6.638	-1.664	0.099
Curvature 1-3	PGMIS	-19.875	7.437	-2.672	0.009

The estimated program differences at spring of first, second, and third grade are little affected by the inclusion of the pretest as a covariate. The new estimates are, respectively, -9.504 ($p = .159$), 13.790 ($p = .064$), and -2.666 ($p = .838$). The school differences are still nonsignificant and the other coefficients show little change.

Summary

The finding of higher immersion strategy achievement on the reading subtest at spring of second grade was confirmed by the analysis of students with pretest scores but the result fell short of statistical significance. As was found for the analysis of all students, the students with pretest scores had higher early-exit achievement in reading at first and third grades but the difference was not statistically significant. The three schools with pretest information exhibited no statistically significant differences in initial status or growth rate.

ERIC

Matching the K-1 and 1-3 Analyses

How well do the K-1 and 1-3 analyses for mathematics match in spring of first grade?

Both the K-1 analyses and the 1-3 analyses produce estimates of the difference between the programs in spring of first grade. The estimates from the K-1 analyses are based on analysis of covariance (ANCOVA), whereas the estimates from the 1-3 analyses are based on hierarchical linear models fitted to individual growth curves using the HLM computer program. If the respective models fit the data reasonably well, the two analytic methods should produce similar estimates in the spring of first grade. Table 98 summarizes some of the estimated program differences for spring of first grade for the mathematics subtest.

Table 98

Two-Program Schools K-1 and 1-3 Analyses:
Estimated Program Differences in Spring 1st for Mathematics

Description	Parameter	Std Err	t-Stat	p-Value
K-1: All Students	-5.201	3.717	-1.399	0.163
K-1: Students with Pretest	3.899	4.992	0.781	0.436
K-1: Both Pretests	5.838	4.755	1.228	0.222
K-1: Pretest Total	5.390	4.876	1.105	0.271
K-1: Students also in 1-3	-6.779	4.099	-1.654	0.100
1-3: All Students	-5.710	3.637	-1.570	0.118
1-3: Students with Pretest	0.619	5.996	0.103	0.918
1-3: Math Pretest	3.853	5.825	0.662	0.510

The estimated program difference is given for five of the K-1 analyses, one from each of the five sets of four analyses. The analysis controlling for schools, absences, and one key covariate was selected as the representative of the set of four analyses. These analyses were labeled "Key Covariates" in the tables of the K-1 analyses. The estimated program difference is also given for all three of the 1-3 analyses presented above.

ERIC

The last of the K-1 analyses listed, labeled "K-1: Students also in 1-3," was run on exactly the same set of students as the first of the 1-3 analyses listed, labeled "1-3: All Students." No other K-1 analysis used the same set of students as any of the 1-3 analyses. The estimated program differences from these two analyses are only about one point apart, much less than the standard error of either estimate. The K-1 analysis on all students also produced a similar estimate of the program difference. These three analyses are gratifyingly well-matched. All three indicate a small early-exit advantage that is not statistically significant.

The K-1 and 1-3 estimates of program differences for students with pretest scores, although run on slightly different groups of students, also match up well. Both the K-1 analyses and the 1-3 analyses produce estimates indicating a small immersion strategy advantage that is not even close to statistical significance. As noted above, this change in sign of the estimated program effect is due primarily to the omission of school 51, where none of the students had pretest scores available.

How well do the K-1 and 1-3 analyses for language match in spring of first grade?

The estimated program differences for the five "Key Covariates" K-1 analyses and the three 1-3 analyses for the language subtest are summarized in Table 99. The two analyses on the same set of students, the "K-1: Students also in 1-3" and "1-3: All Students" analyses, produce estimates of 8.432 and 11.037 for the program difference in spring of first grade. While the difference between these estimates is only 2.605, about half of the standard error of either estimate, this difference is enough to make the 1-3 estimate statistically significant ($p = .037$) even though the K-1 estimate is not ($p = .104$).

ERIC

Table 99

Two-Program Schools K-1 and 1-3 Analyses:
Estimated Program Differences in Spring 1st for Language

Description	Parameter	Std Err	t-Stat	p-Value
K-1: All Students	5.004	4.638	1.079	0.282
K-1: Students with Pretest	7.984	6.464	1.235	0.219
K-1: Both Pretests	9.757	6.350	1.537	0.127
K-1: Pretest Total	9.920	6.342	1.564	0.120
K-1: Students also in 1-3	8.432	5.156	1.635	0.104
1-3: All Students	11.037	5.245	2.104	0.037
1-3: Students with Pretest	17.645	7.484	2.358	0.020
1-3: Pretest Total	20.317	7.488	2.713	0.008

The K-1 and 1-3 estimates of program differences for students with pretest scores are not as close, with the 1-3 estimate between one and two standard errors larger than the K-1 estimate. Again the estimates from the K-1 analyses are not statistically significant (p greater than .10), but the estimates from the 1-3 analyses are significant (p = .020 and p = .008).

Overall, the K-1 and 1-3 estimates of the difference between the two programs are fairly consistent, all indicating an advantage on the language subtest for the immersion strategy students. The 1-3 analyses, which make use of more data than the K-1 analyses, are probably slightly more accurate. It seems reasonable to conclude that the immersion strategy students show about a five to fifteen point advantage in language at spring of first grade.

<u>How well do the K-1 and 1-3 analyses for reading match in spring of first grade?</u>

The estimated program differences on the reading subtest for the five "Key Covariates" K-1 analyses and the three 1-3 analyses are given in Table 100. The two analyses on the same set of students produce estimates of -4.784 (for the K-1 analysis on students also in the 1-3 analysis) and

-4.442 (for the 1-3 analysis on "All Students"), which are remarkably close. Both estimates indicate a small early-exit advantage, but neither is statistically significant (p = .260 and p = .320, respectively).

Table 100

Two-Program Schools K-1 and 1-3 Analyses:
Estimated Program Differences in Spring 1st for Reading

Description	Parameter	Std Err	t-Stat	p-Value
K-1: All Students	-8.177	3.773	-2.167	0.031
K-1: Students with Pretest	-14.619	5.232	-2.794	0.006
K-1: Both Pretests	-13.707	5.226	-2.623	0.010
K-1: Pretest Total	-13.586	5.216	-2.605	0.010
K-1: Students also in 1-3	-4.784	4.233	-1.130	0.260
1-3: All Students	-4.442	4.45*	-0.998	0.320
1-3: Students with Pretest	-11.264	6.71*	-1.679	0.096
1-3: Pretest Total	-9.504	6.75*	-1.407	0.163

* Estimated from t-statistic

Note that the estimated first grade program difference for the 1-3 analyses comes from combining the program effect for predicting initial status and the program effect for the curvature. Because this estimate is a combination of two coefficients, the HLM program does not report a standard error directly. However, the hypothesis testing feature of HLM permits calculation of an approximate t-statistic from which an estimated standard error can be calculated.

The K-1 estimate of program differences for all students in the K-1 analyses is -8.177, which is significant at the .05 level (p = .031), even though it is less than one standard error from the nonsignificant estimate obtained using students also in the 1-3 analyses. In the K-1 analyses, the students with pretest scores (and therefore not including school 51) show an even larger early-exit advantage in spring of first grade, with the differences all significant at the .01 level. The 1-3 analyses also indicate a relatively larger early-exit advantage in English reading at spring of first grade for the students with a pretest. Although the estimated program differences from the 1-3 analyses on students with pretest scores are within one standard error of the corresponding K-1 estimates, the 1-3 estimates are not statistically significant.

Overall, the K-1 and 1-3 estimates of the difference between the two programs have the same sign, all indicating an advantage on the English reading subtest for the early-exit students. The magnitude of the difference is very close for the two analyses on exactly the same students and is not significant. The other analyses show larger differences, with the differences achieving significance for the K-1 analyses but not for the 1-3 analyses. Thus the program difference in reading in the spring of first grade is sensitive to the exact subset of students used in the analysis. No firm conclusion is possible; however, if the early-exit students at the end of first grade are higher in reading achievement than immersion strategy students, the difference is likely to be no more than ten points.

Conclusions

What are the overall conclusions from the two-program schools analyses?

The results of the two-program schools analyses differ by outcome measure (see Table 101 for a summary of results). The mathematics subtest shows essentially no differences between the programs. For the language subtest, the students in immersion strategy programs have a higher status at spring of first grade, with the difference marginally statistically significant. While the students in immersion strategy programs have a slightly lower growth rate than students in early-exit programs, the difference is not statistically significant. Taken together, the estimated difference in first grade scores and the difference in 1-3 growth rates cancel each other out by spring of third grade: students in the two programs are not significantly different as of the testing in spring of third grade.

Table 101

Summary of Program Differences in Two-Program School HLM Analyses

		Initial Status	Growth Rate	Curvature
Mathematics (see Table 89)	ESS Differences+	-5.710	1.428	-6.959
	p-Value	.118	.606	.000**
English Language (see Table 90)	ESS Differences	11.037	-4.304	-10.253
	p-Value	.037*	.377	.002
English Reading (see Table 91)	ESS Differences	1.406	-1.689	-17.543
	p-Value	.745	.676	.002*

\+ Expanded Scale Score Differences between IS and EE
* Statistically significant at p less than .05
** Statistically significant at p less than .01

For the English reading subtest, early-exit students have higher scores in spring of first grade, but the exact amount of the difference and its statistical significance is sensitive to the form of the analysis and to which subset of the students is included. Because of the consistency of the sign and magnitude of the difference, however, one can with some confidence conclude that the early-exit students have a slight edge in reading achievement in spring of first grade. As noted earlier, this is consistent with the hypothesis that teaching students to read in their first language allows them to develop skills which they may apply when learning English skills. Perhaps the most surprising finding is the program difference in curvature of the growth curves from grades one through three. This difference results in higher scores for immersion strategy students in spring of second grade. However, by spring of third grade the early-exit students once again have higher scores than the immersion strategy students, but the difference is not statistically significant.

A similarity between the results for the language and reading subtests is that on both subtests the immersion strategy students show an early but temporary boost in achievement. A difference between the two tests is that the boost appears at spring of first grade for the language subtest and at spring of second grade for the reading subtest. While it is tempting to conclude that the greater growth rates for early-exit students after first grade (for language) or second grade (for reading) suggest the beginning of an upward trend by early-exit students, it is not appropriate to extrapolate the growth curves beyond spring of third grade.

ERIC

IV. THE RELATIVE EFFECTIVENESS OF IMMERSION STRATEGY AND EARLY-EXIT TRANSITIONAL BILINGUAL EDUCATION PROGRAMS IN ONE-PROGRAM SCHOOLS

Introduction

What was the purpose of the one-program schools analyses?

The analyses of one-program schools were designed to determine the relative effectiveness of the immersion strategy (IS) and early-exit (EE) programs controlling for differences between schools in the analyses. Program effectiveness was evaluated in the same three areas of academic achievement that were evaluated in the two-program schools analyses: mathematics, language arts, and reading, as measured in English.

Why were the one-program schools analyzed separately?

For the two-program schools, it was possible to compare the programs within each of the schools, thereby controlling for school-level effects. For the one-program schools, the programs can be evaluated by comparing the IS schools with the EE schools. It would have been possible simply to ignore school effects and relegate school-level variation to unexplained between-student variation. Under certain circumstances, this approach would be justified. For these analyses, however, school-level variation was found to be too large to ignore.

Were the IS and EE students comparable in the one-program schools?

Analyses of the students in the one-program schools revealed many statistically significant differences between IS and EE students beyond those directly related to the program. That is, there were substantial "selection effects" in the one-program schools: when they entered the program in kindergarten, the students in IS schools were materially different from the students in EE schools, even after controlling for district. No doubt many of these differences were due to the differences in the neighborhoods served by the schools.

The presence of selection effects means that the two groups of students are not comparable. Any difference in the achievement for the IS and EE students might be attributable to the differences between the two groups of students. It is possible to adjust the estimated program difference for these differences by including a "propensity score" that expresses the tendency for a student to be in one program or the other. The analyses in this chapter include a propensity score to adjust for the detected differences between IS students and EE students. If the IS and EE students were completely dissimilar on some characteristic that was not related to program (e.g., height), the propensity score would not allow us to disentangle any program differences in achievement. The more similar the two groups of students, the less impact the propensity score has on the estimated program effect. For additional details on the calculation and use of propensity scores, see Chapter II.

What statistical methods were used to compare the two programs?

The one-program schools analyses were carried out in the same two steps as the two-program schools analyses: a set of K-1 analyses and a set of 1-3 analyses. The K-1 analyses were used to assess the achievement of each child at the spring of first grade. Background variables and, when available, the fall kindergarten test scores were used as predictors of the spring first grade test scores for mathematics, language arts, and reading. Analysis of Covariance (ANCOVA) was used to perform the K-1 analyses. Because it was desired to allow random school effects, the HLM computer program was used to perform the ANCOVA. The HLM computer program allows us to specify school effects as random (Bryk et al., 1988).

The main set of 1-3 analyses was parallel to the two-program schools analyses. The HLM program was used to develop a two-level hierarchical linear model: one model representing an individual growth curve for each child and another model relating those growth curves to student attributes.

The same form of the CTBS test scores (the Expanded Scale Scores, or ESS) and the same measure of time (zero at spring of first grade and increasing by one every grade thereafter) were used for the one-program schools analyses as were used for the two-program schools analyses.

Additional 1-3 analyses were performed using the three-level HLM program. The three-level HLM program, described in more detail in Chapter II, permits both student-level random effects (individual student differences in slope and intercept) and school-level random effects (school differences in slope and intercept). That is, an individual growth curve is generated for each student and independently for each school. Both the individual student and the school growth curves are used to generate overall growth curves for groups of students (e.g., between programs, between schools, etc.). Using the three-level HLM program permits the 1-3 analyses of one-program schools to be more similar to the K-1 analyses, where schools are treated as random. Treating schools as random is a preferable analytic approach, and it would have been desirable to use the three-level HLM program for all the 1-3 analyses for the one-program schools. However, the program is still under development, so it was only possible to use it to confirm some of the main set of 1-3 analyses.

<u>Were the schools significantly different from each other?</u>

Even after including the background variables in the model, statistically significant school effects were found. School membership was found to be a significant predictor of both initial status and growth rate, just as it was in the two-program schools analyses. Although it is possible the school differences could be explained by aspects of the program implementation in the schools, such analyses go beyond the scope of nominal program analyses.

What are the implications of school differences among the one-program schools?

The presence of differences among the one-program schools greatly complicates the comparison of the immersion strategy and early-exit programs. The one-program school analyses cannot control for school effects because school is confounded with program. As a result, we cannot directly separate out school effects from program effects. The differences between the programs must be considered in light of the school differences: the programs are compared by assessing whether the difference between immersion strategy schools and early-exit schools is large compared to the variability among schools within programs.

District differences must be assessed in the same way: by comparing differences between districts to the variability among schools within the districts. With only 29 one-program schools, it is difficult to evaluate district differences. Most of the schools are in district H/I, but the number of students in each school is relatively small. It is therefore difficult to evaluate the amount of within-program, within-district variability of schools. Only some of the district-program combinations can contribute to the estimate of that variability, because some districts have only one IS school or only one EE school.

In an extensive series of analyses not presented in detail here, the school variability within district-program combination was evaluated and compared with the district variability within program. The school variability was found to be highly statistically significant and therefore not ignorable. Furthermore, the district variability within program was found to be only slightly greater than the school variability within district and program. That is, districts varied within program as much as schools varied within district and program. As a result, the variability between districts within a program could simply be a reflection of the variability between schools within districts. Although the districts differed, the differences were not much greater than the amount expected in light of the school variability.

Why were schools combined into groups?

The fact that schools exhibited considerable variation even within district and program meant that schools had to be included in the models. Some schools had too few students to permit reliable estimates of school effects. For this reason, school groups were created.

Schools were never combined across the two programs or across the five districts. In fact, only schools in district C and district H were combined; all other schools had at least nine students and combining the schools was deemed unnecessary. In district C, the two immersion strategy schools were combined and the two early-exit schools were combined. However, a test needed to be done to ensure that there were no differences in the growth curves of the schools that were combined. Preliminary HLM modeling not presented here had revealed no statistically significant differences in the growth curves of the students within the pairs of schools. This ensures that schools were combined only if the students within a pair of schools had the same growth curves so as not to mask true growth curve differences that might exist between schools. With only five students in one of the two IS schools and four students in one of the two EE schools, separate school estimates would have been unstable.

In district H, the immersion strategy district paired with early-exit district I, the twelve schools had from three to twenty students each. (As a reminder, while districts H and I have different programs, they are in the northeast United States and are contiguous districts. They serve essentially the same population. The similarities are such that the two districts are considered appropriate for combination.) HLM analyses not presented here indicated that there were statistically significant school differences within district H, so combining all the schools into one group was considered inappropriate. Instead, the schools were combined into four groups: one group of six schools (with 39 students), one group of three schools (with 31 students), one group of two schools (with 13 students), and the single remaining school (with 20 students). The school groups are indicated in Table 14 in Chapter II. The school combinations in district H were arrived at empirically. That is, the HLM growth curves for the students in each school or school group were compared and the two schools or school groups with most similar growth curves were combined. This iterative process was continued until a small set of relatively homogeneous groups was obtained.

This process of combining groups of schools based on growth curve results introduces a bias into the analytic results. Using empirically-created groups tends to decrease the within-group variability and increase the between-group variability. This can produce the appearance of school differences where none actually exist.

Because of this empirical combination of schools in district H, the district H school groups are likely to exhibit greater differences from each other than individual schools in the district. This bias should be considered when evaluating the results of the analyses. It must be emphasized, however, that the original twelve schools in district H were found to be significantly different from each other (even with only a handful of students in many of the schools). Furthermore, the other districts also exhibited significant school-to-school variability within program. Thus the overall finding of significant school variability is not dependent upon the use of empirical results to form school groups in district H.

<u>What is the pattern of schools and school groups across districts and programs</u>?

Both districts A and B have one immersion strategy school and three early-exit schools. District C has two schools in each program, but these were combined into a single school group in each program, as explained above. District H has twelve immersion strategy schools combined into four school groups, and district I has four early-exit schools; these nearby districts are treated as a single district in the analyses. District F has only one early-exit school and no immersion strategy schools. When district is controlled, the district F school does not contribute to the estimated program effect because there is no comparison school in that district.

How did the K-1 analyses accommodate school differences?

The K-1 analyses for the one-program schools parallel those for the two-program schools in Chapter III, but the ANCOVA analyses were performed using the HLM computer program with schools treated as random effects. The difference between students in immersion strategy schools and students in early-exit schools was evaluated at spring of first grade for each of the three subtests. The K-1 analyses were performed once with each of the 29 individual schools treated separately and then again with the schools combined into the 19 groups. The results for the two sets of analyses were similar. For ease of comparison with the 1-3 analyses, only the analyses for the 19 groups will be presented.

How did the 1-3 analyses accommodate school differences?

The main set of 1-3 analyses used the two-level HLM program with schools treated as fixed effects. (This set of analyses parallels the analyses in Chapter III for two-program schools.) That is, the schools and school groups were represented in the analysis by a set of variables. School membership was used to predict the initial status and growth rate of the students in that school. An overall program difference is estimated essentially by comparing the average immersion strategy school with the average early-exit school. Unfortunately, with the fixed-effects model the standard error calculated for the estimated program difference is too small because it does not reflect the school-to-school variability. If no significant program difference is found, this is not critical. If the program difference is large compared to its standard error, however, this apparent statistical significance cannot be formally evaluated.

The three-level program permits the schools to be treated as random effects. This allows us to properly reflect the school-to-school variability. This analysis was done as an extra effort to verify whether

ERIC

it was necessary to do the HLM analyses with school as a random effect. If the results of the HLM analysis with school as a random effect were consistent with the analysis where school was treated as a fixed effect, this would increase the analytic power behind the findings. School membership is still used to predict the initial status and growth rate of the students in that school, but the individual schools and school groups do not appear explicitly in the model. However, as noted above, the variability among the schools within each program is used as an estimate of the amount of school variability to be expected within each program. The programs are then compared in light of the school-level variability. This results in a more accurate standard error for the program effect.

<u>**How were districts accommodated in the K-1 and 1-3 analyses**</u>?

For the two-program schools, school-level variation can be controlled by comparing the programs within each school. In a similar way, district-level variation can be controlled in the one-program schools analyses by comparing the programs within each district. As noted above, with so few schools it is difficult to assess district differences. Even sizable differences among the districts may not reach statistical significance.

For the K-1 analyses, all of the alternative models were estimated with and without including district variables. This permits the comparison of results assuming that there are district differences and assuming that there are no district differences. The results were generally very similar and the district differences usually were not statistically significant.

For the main set of 1-3 analyses, using the two-level HLM program, district differences were not estimated. Each individual school or school group appears in the model. District differences can be inferred by comparing the schools in one district with the schools in another district. While multivariate contrasts can be used to make these comparisons, they are not reported here. Instead, the most important 1-3 analyses were repeated using the three-level HLM program. The three-level program was used to confirm the estimated difference between the two programs. It was also used to assess district effects and district-program interactions.

Results of the K-1 Analyses for Mathematics

The following section details the K-1 HLM analyses for mathematics. Please see Appendix C for the average first grade mathematics scores, unadjusted and adjusted, by program. They document the similarity between immersion strategy and early-exit first grade students.

What were the results of the basic K-1 analyses for the mathematics subtest?

The basic K-1 analyses used ANCOVA with school as a random effect to evaluate program differences in achievement at spring of first grade. The two-level HLM computer program was used to perform the ANCOVA. Analyses were performed with no covariates; with the propensity score as a covariate, to adjust for initial differences between the students in the two programs; with the propensity score, the number of absences in first grade, and two "key covariates" that varied according to the subtest; and with the propensity score, the number of absences and five background variables found to be significant predictors of achievement. The set of five background variables included in these one-program schools analyses are the same five included in the corresponding two-program analyses: EDAVG (the average years of education of the student's parents), PRESCHY (whether the student attended preschool), FEMALE (whether the student was a girl), BOOKSHM (the number of books in the student's home), and ANYEPTOP (whether the parents reported using English when talking to each other). The analyses were also performed with and without adjusting for district effects.

Table 102 summarizes the results of the analysis of program differences for the mathematics subtest not controlling for any other variable. It shows an estimated BASE (overall average) and program difference (PGMIS) at spring of first grade. The estimated coefficient for PGMIS is almost exactly zero, indicating no difference between the two programs. It should be reemphasized that the program difference is based on only 19 schools and school groups. Because each school has only one program, program differences are assessed by comparing the distribution of school averages in the immersion strategy program with the distribution of school averages in the early-exit program. Thus the approximate t-statistics in this table have only 17 degrees of freedom, and relatively large values of the t-statistics will be needed to achieve statistical significance compared with the tables for the two-program schools analyses.

Table 102

One-Program Schools K-1 Analysis for Mathematics
All Students; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	250.916	4.954	50.653	0.000
Spring 1st	PGMIS	-0.016	8.130	-0.002	0.998

Table 103 summarizes the results of the analysis of program differences after including PONEK1, the propensity score developed for the students in these one-program schools K-1 analyses. Table 104 adds three "key" covariates to the propensity score: ABS1 (the number of absences in first grade), EDAVG, and FEMALE. The other three background variables were not close to statistical significance for the mathematics subtest. Table 105 includes all the covariates: the propensity score, the number of absences, and the five background variables. Although the program effect increases slightly as more covariates are added, it never even approaches statistical significance.

Table 103

One-Program Schools K-1 Analysis for Mathematics
All Students; Program and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	253.933	6.669	38.075	0.000
Spring 1st	PGMIS	0.721	8.618	0.084	0.934
Spring 1st	PONEK1	-6.994	9.672	-0.723	0.480

Table 104

One-Program Schools K-1 Analysis for Mathematics
All Students; Program and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	243.837	8.119	30.032	0.000
Spring 1st	PGMIS	1.208	7.222	0.167	0.870
Spring 1st	PONEK1	7.126	9.912	0.719	0.485
Spring 1st	ABS1	-0.517	0.169	-3.052	0.009
Spring 1st	EDAVG	1.375	0.438	3.136	0.008
Spring 1st	FEMALE	-6.466	2.564	-2.522	0.025

Table 105

One-Program Schools K-1 Analysis for Mathematics
All Students; Program and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	241.348	8.089	29.838	0.000
Spring 1st	PGMIS	1.344	6.999	0.192	0.852
Spring 1st	PONEK1	11.090	10.683	1.038	0.324
Spring 1st	ABS1	-0.521	0.170	-3.070	0.012
Spring 1st	EDAVG	1.320	0.470	2.809	0.019
Spring 1st	PRESCHY	-2.847	3.030	-0.940	0.370
Spring 1st	FEMALE	-6.941	2.581	-2.690	0.023
Spring 1st	BOOKSHM	1.475	1.000	1.475	0.171
Spring 1st	ANYEPTOP	-2.343	3.345	-0.701	0.500

Table 102 through Table 105 do not control for any possible district effects. By including district dummy variables as predictors, the estimated program differences can be controlled for district differences. Essentially, the early-exit schools in each district are compared with the immersion strategy schools in the same district to estimate the program difference. Controlling for district in these one-program schools analyses corresponds to controlling for school in the two-program schools analyses. However, there are only a few schools in each district, so district differences are hard to distinguish statistically from school differences. After adding four district dummy variables to the BASE and PGMIS variables, only 13 degrees of freedom are left for school-to-school variation among the 19 schools.

Table 106 through Table 109 show the results after controlling for district. The district variables are coded as zero-one dummy variables, so that the omitted district (district C) serves as a "reference" district. The parameter estimate for BASE is the estimate for district C. The parameter estimates for the other four district variables (DISTA, DISTB, DISTF, and DISTHI) represent the difference between that district and district C.

Table 106

One-Program Schools K-1 Analysis for Mathematics
All Students; Program and District

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	251.146	11.978	20.967	0.000
Spring 1st	PGMIS	-1.560	8.224	-0.190	0.852
Spring 1st	DISTA	4.263	13.997	0.305	0.766
Spring 1st	DISTB	-15.805	14.158	-1.116	0.284
Spring 1st	DISTF	14.770	20.273	0.729	0.479
Spring 1st	DISTHI	4.621	12.707	0.364	0.722

Table 107

One-Program Schools K-1 Analysis for Mathematics
All Students; Program, District, and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	259.541	12.378	20.968	0.000
Spring 1st	PGMIS	0.282	8.057	0.035	0.973
Spring 1st	DISTA	4.988	13.608	0.367	0.720
Spring 1st	DISTB	-21.709	14.081	-1.542	0.149
Spring 1st	DISTF	28.559	20.897	1.367	0.197
Spring 1st	DISTHI	8.713	12.523	0.696	0.500
Spring 1st	PONEK1	-22.183	11.130	-1.993	0.069

Table 108

One-Program Schools K-1 Analysis for Mathematics
All Students; Program, District, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	255.846	13.292	19.248	0.000
Spring 1st	PGMIS	0.813	7.640	0.106	0.918
Spring 1st	DISTA	1.916	12.915	0.148	0.885
Spring 1st	DISTB	-18.589	13.465	-1.381	0.201
Spring 1st	DISTF	18.696	20.331	0.920	0.382
Spring 1st	DISTHI	5.201	12.127	0.429	0.678
Spring 1st	PONEK1	-10.582	12.901	-0.820	0.433
Spring 1st	ABS1	-0.518	0.170	-3.049	0.014
Spring 1st	EDAVG	0.975	0.477	2.044	0.071
Spring 1st	FEMALE	-6.547	2.561	-2.556	0.031

ERIC

Table 109

One-Program Schools K-1 Analysis for Mathematics
All Students; Program, District, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	254.281	13.420	18.948	0.000
Spring 1st	PGMIS	1.063	7.519	0.141	0.892
Spring 1st	DISTA	1.252	12.744	0.098	0.925
Spring 1st	DISTB	-18.391	13.438	-1.369	0.220
Spring 1st	DISTF	18.505	20.111	0.920	0.393
Spring 1st	DISTHI	4.584	11.934	0.384	0.714
Spring 1st	PONEK1	-8.940	14.306	-0.625	0.555
Spring 1st	ABS1	-0.516	0.170	-3.035	0.023
Spring 1st	EDAVG	0.877	0.517	1.697	0.141
Spring 1st	PRESCHY	-1.334	3.104	-0.430	0.682
Spring 1st	FEMALE	-7.026	2.578	-2.725	0.034
Spring 1st	BOOKSHM	1.558	1.002	1.555	0.171
Spring 1st	ANYEPTOP	-2.261	3.361	-0.672	0.526

Even after adjusting for district, the program differences in all four tables are very small and nonsignificant. Although the estimated district differences are much larger than the estimated program difference, the test for overall district differences is nonsignificant at the .05 level for all four tables. Only for the analysis controlling for propensity but no other covariates (see Table 107) does the overall test approach significance (p = .053). The near-significant district differences in this analyses are eliminated by the addition of the key covariates (see Table 108). To the extent there are any district differences, district F is high and district B is low in mathematics achievement as measured in English at spring of first grade.

The absence of statistically significant district differences means that, within each program, schools vary about as much within a district as they do across the districts. However, with so few districts having multiple schools or school groups with the same program, the statistical test for district effects does not have much statistical power. That means that there may be district differences, but with only a few schools to compare the differences may not achieve statistical significance.

Summary

These one-program schools show no statistically significant program or district differences on the mathematics subtest at spring of first grade. Program differences are very small regardless of the covariates included as predictors. District differences are much larger in magnitude but never achieve statistical significance at the .05 level. The addition of key covariates tends to reduce the size of district differences.

How were the K-1 analyses for mathematics affected by controlling for pretest?

The basic K-1 analyses for mathematics estimate program effects with and without controlling for propensity, district, and several covariates. It is possible that the results would be changed by controlling for pretest factors. As was done for the two-program schools, the evaluation of the pretest effect is accomplished in two parts. First, it is necessary to select students who have a pretest, and second, by adding the pretest as a covariate for this reduced set of students.

Reducing the students analyzed to those who have a pretest means leaving out three of the 19 school groups. One IS school group is eliminated from district H, and one EE school is eliminated from district I. Also, the only school in district F is eliminated. (The eliminated school is in a very rural site in the southwest United States. This school serves all kindergarten through third grade students. With the loss of funding for its immersion strategy program, the district began phasing out the immersion strategy program beginning with first grade. As a result, the original kindergarten cohort was dropped from the study as there was no program for the students in which to continue.) Table 110 summarizes the unadjusted program difference on the mathematics subtest for the students with a pretest. Table 111 shows the analysis after including the propensity score. Table 112 and Table 113 reflect the addition of the three key covariates and all covariates, respectively. The estimated program differences are still extremely small.

Table 110

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	248.657	5.442	45.694	0.000
Spring 1st	PGMIS	-0.109	8.861	-0.012	0.990

Table 111

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	251.238	7.029	35.745	0.000
Spring 1st	PGMIS	0.938	9.391	0.100	0.922
Spring 1st	PONEK1	-7.237	11.926	-0.607	0.554

Table 112

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	247.615	9.223	26.848	0.000
Spring 1st	PGMIS	0.626	8.449	0.074	0.942
Spring 1st	PONEK1	2.018	12.428	0.162	0.874
Spring 1st	ABS1	-0.459	0.190	-2.420	0.036
Spring 1st	EDAVG	0.958	0.541	1.772	0.107
Spring 1st	FEMALE	-7.757	3.050	-2.544	0.029

Table 113

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	246.318	9.331	26.398	0.000
Spring 1st	PGMIS	0.787	8.485	0.093	0.929
Spring 1st	PONEK1	1.092	13.673	0.080	0.939
Spring 1st	ABS1	-0.460	0.190	-2.425	0.046
Spring 1st	EDAVG	0.864	0.597	1.447	0.191
Spring 1st	PRESCHY	-0.162	3.601	-0.045	0.965
Spring 1st	FEMALE	-8.194	3.055	-2.682	0.031
Spring 1st	BOOKSHM	1.974	1.185	1.666	0.140
Spring 1st	ANYEPTOP	-6.507	4.189	-1.553	0.164

Table 114, Table 115, Table 116, and Table 117 show the results of adding district as a predictor. The estimated program differences remain very small and are much smaller than the estimated district differences. Although district B continues to have markedly lower mathematics achievement than the other districts, none of the differences are statistically significant. The analysis for all students controlling for district and propensity (see Table 107) showed district effects that were near significance (p = .053). The corresponding table for students with a pretest, which necessarily omits district F, shows similar district differences among the three remaining districts, and again the difference is not significant (p = .092). The district differences are little changed by including background variables and do not approach statistical significance.

Table 114

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and District

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	253.284	11.915	21.257	0.000
Spring 1st	PGMIS	-2.578	8.762	-0.294	0.774
Spring 1st	DISTA	4.376	14.021	0.312	0.761
Spring 1st	DISTB	-18.177	13.957	-1.302	0.219
Spring 1st	DISTHI	-0.365	13.041	-0.028	0.978

Table 115

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, District, and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	260.746	12.817	20.344	0.000
Spring 1st	PGMIS	-1.048	8.732	-0.120	0.907
Spring 1st	DISTA	4.762	13.871	0.343	0.738
Spring 1st	DISTB	-23.413	14.251	-1.643	0.131
Spring 1st	DISTHI	2.792	13.076	0.213	0.835
Spring 1st	PONEK1	-19.584	13.229	-1.480	0.170

Table 116

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, District, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	263.170	14.345	18.345	0.000
Spring 1st	PGMIS	-0.865	8.330	-0.104	0.920
Spring 1st	DISTA	2.134	13.303	0.160	0.877
Spring 1st	DISTB	-22.208	13.744	-1.616	0.150
Spring 1st	DISTHI	1.734	12.894	0.134	0.897
Spring 1st	PONEK1	-14.584	15.222	-0.958	0.370
Spring 1st	ABS1	-0.463	0.190	-2.434	0.045
Spring 1st	ELAVG	0.521	0.595	0.876	0.410
Spring 1st	FEMALE	-7.857	3.048	-2.578	0.037

Table 117

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, District, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	263.621	14.613	18.041	0.000
Spring 1st	PGMIS	-0.593	8.212	-0.072	0.946
Spring 1st	DISTA	2.200	13.164	0.167	0.875
Spring 1st	DISTB	-24.153	13.842	-1.745	0.156
Spring 1st	DISTHI	0.765	12.721	0.060	0.955
Spring 1st	PONEK1	-18.970	16.992	-1.116	0.327
Spring 1st	ABS1	-0.457	0.190	-2.410	0.074
Spring 1st	EDAVG	0.334	0.662	0.505	0.640
Spring 1st	PRESCHY	1.560	3.677	0.424	0.693
Spring 1st	FEMALE	-8.237	3.052	-2.699	0.054
Spring 1st	BOOKSHM	2.180	1.187	1.836	0.140
Spring 1st	ANYEPTOP	-6.268	4.213	-1.488	0.211

Table 118 through Table 121 show the effect of adding the mathematics pretest (ANALPREM) as a covariate in the analyses that do not control for district. The mathematics pretest is highly significant as a predictor of first grade mathematics achievement. The language pretest was not significant as an additional predictor, so it is not included in the analyses. The mathematics pretest alone was found to be a better predictor than the sum of the mathematics and language subtests. The estimated program effects are still small and nonsignificant, regardless of the other covariates included.

Table 118

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	150.749	15.273	9.870	0.000
Spring 1st	PGMIS	2.126	7.831	0.271	0.790
Spring 1st	ANALPREM	0.577	0.085	6.747	0.000

Table 119

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, Propensity, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	149.553	17.598	8.498	0.000
Spring 1st	PGMIS	1.902	8.019	0.237	0.816
Spring 1st	PONEK1PP	1.602	11.515	0.139	0.892
Spring 1st	ANALPREM	0.581	0.090	6.418	0.000

Table 120

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, Key Covariates, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	146.663	18.692	7.846	0.000
Spring 1st	PGMIS	1.802	7.428	0.243	0.814
Spring 1st	PONEK1PP	8.399	11.741	0.715	0.493
Spring 1st	ABS1	-0.331	0.181	-1.827	0.101
Spring 1st	EDAVG	0.951	0.497	1.913	0.088
Spring 1st	FEMALE	-7.577	2.889	-2.623	0.028
Spring 1st	ANALPREM	0.576	0.090	6.416	0.000

Table 121

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, All Covariates, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	146.688	18.872	7.773	0.000
Spring 1st	PGMIS	1.862	7.345	0.253	0.808
Spring 1st	PONEK1PP	9.046	12.815	0.706	0.507
Spring 1st	ABS1	-0.337	0.181	-1.858	0.113
Spring 1st	EDAVG	0.870	0.549	1.585	0.164
Spring 1st	PRESCHY	-1.086	3.393	-0.320	0.760
Spring 1st	FEMALE	-8.002	2.900	-2.759	0.033
Spring 1st	BOOKSHM	1.695	1.126	1.506	0.183
Spring 1st	ANYEPTOP	-4.776	3.990	-1.197	0.276
Spring 1st	ANALPREM	0.568	0.091	6.254	0.001

Table 122 through Table 125 show the effect of adding the mathematics pretest as a covariate in the analyses that do not control for district. The estimated program effects are still small and nonsignificant regardless of the other covariates included. In addition, the tests for district differences remain nonsignificant.

Table 122

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, District, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	157.365	18.522	8.496	0.000
Spring 1st	PGMIS	-0.314	8.333	-0.038	0.971
Spring 1st	DISTA	-0.177	13.344	-0.013	0.990
Spring 1st	DISTB	-13.985	13.280	-1.053	0.317
Spring 1st	DISTHI	0.285	12.394	0.023	0.982
Spring 1st	ANALPREM	0.564	0.086	6.545	0.000

Table 123

One-Program Schools K-1 Analysis for Mathematics
Students with Pretest; Program, District,
Propensity, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	163.920	21.287	7.700	0.000
Spring 1st	PGMIS	0.300	8.352	0.036	0.972
Spring 1st	DISTA	0.479	13.319	0.036	0.972
Spring 1st	DISTB	-15.907	13.568	-1.172	0.271
Spring 1st	DISTHI	1.533	12.495	0.123	0.905
Spring 1st	PONEK1PP	-8.094	12.996	-0.623	0.549
Spring 1st	ANALPREM	0.541	0.094	5.772	0.000

Table 124

One-Program Schools K-1 Analysis for Mathematics Students with Pretest; Program, District, Key Covariates, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	163.812	22.874	7.161	0.000
Spring 1st	PGMIS	0.453	7.953	0.057	0.956
Spring 1st	DISTA	-2.538	12.786	-0.199	0.849
Spring 1st	DISTB	-14.686	12.997	-1.130	0.302
Spring 1st	DISTHI	-0.558	12.224	-0.046	0.965
Spring 1st	PONEK1PP	-1.572	14.348	-0.110	0.916
Spring 1st	ABS1	-0.335	0.182	-1.847	0.114
Spring 1st	EDAVG	0.717	0.541	1.326	0.233
Spring 1st	FEMALE	-7.684	2.894	-2.655	0.038
Spring 1st	ANALPREM	0.540	0.095	5.696	0.001

Table 125

One-Program Schools K-1 Analysis for Mathematics Students with Pretest; Program, District, All Covariates, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	165.878	23.493	7.061	0.006
Spring 1st	PGMIS	0.626	7.798	0.080	0.941
Spring 1st	DISTA	-2.420	12.571	-0.193	0.860
Spring 1st	DISTB	-15.609	12.966	-1.204	0.315
Spring 1st	DISTHI	-1.302	11.999	-0.109	0.920
Spring 1st	PONEK1PP	-3.024	15.936	-0.190	0.862
Spring 1st	ABS1	-0.338	0.182	-1.860	0.160
Spring 1st	EDAVG	0.584	0.601	0.971	0.403
Spring 1st	PRESCHY	-0.090	3.470	-0.026	0.981
Spring 1st	FEMALE	-8.077	2.905	-2.781	0.069
Spring 1st	BOOKSHM	1.828	1.132	1.615	0.205
Spring 1st	ANYEPTOP	-4.539	4.026	-1.127	0.342
Spring 1st	ANALPREM	0.525	0.097	5.415	0.012

Summary

For the mathematics subtest, reducing the sample to students with pretest scores has little effect. Adding the mathematics pretest, even though it has considerable predictive power, does not change the fundamental finding of no program differences. Although district B consistently has the lowest mathematics test scores, there are no statistically significant differences among the districts. That is, the districts remain in the same relationship to each other.

What were the conclusions from the K-1 analyses for the mathematics subtest?

There are no detectable differences in mathematics achievement in spring of first grade between students in immersion strategy schools and students in early-exit schools. The estimated program difference never reaches three ESS points and never approaches statistical significance. Although the estimated district differences in mathematics achievement are much larger than the estimated program difference, they are never statistically significant. All of these K-1 analyses for the mathematics subtest are insensitive to the covariates included. This finding is consistent with the hypothesized similarity in achievement between immersion strategy and early-exit students in light of the fact that early-exit teachers tended not to provide content instruction in their students' primary language, but taught in English employing essentially the same teaching strategies.

Results of the K-1 Analyses for Language

What were the results of the basic K-1 analyses for the language subtest?

Table 126 shows the effect of program for the language subtest without controlling for district or any covariates, and Table 127 shows the results after controlling for the propensity score. Table 128 adds the three "key covariates" for the language subtest: ABS1 (the number of absences in first grade), BOOKSHM (the number of books in the home), and ANYEPTOP (whether English is used between the parents). Even though ANYEPTOP is not statistically significant for the language subtest, both BOOKSHM and ANYEPTOP were included as key covariates for ease of comparison with the results for the reading subtest. Table 129 shows the effect of adding the other three background variables as covariates. The estimated program difference ranges from an IS advantage of about three points to a little over four points, bu the difference is never statistically significant.

Table 126

One-Program Schools K-1 Analysis for Language
All Students; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	253.295	7.212	35.119	0.000
Spring 1st	PGMIS	4.311	11.841	0.364	0.720

Table 127

One-Program Schools K-1 Analysis for Language
All Students; Program and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	247.857	9.034	27.437	0.000
Spring 1st	PGMIS	2.946	11.479	0.257	0.801
Spring 1st	PONEK1	12.661	13.450	0.941	0.361

ERIC

Table 128

One-Program Schools K-1 Analysis for Language
All Students; Program and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	243.391	9.216	26.411	0.000
Spring 1st	PGMIS	3.507	10.351	0.339	0.740
Spring 1st	PONEK1	13.649	13.060	1.045	0.315
Spring 1st	ABS1	−0.412	0.242	−1.704	0.112
Spring 1st	BOOKSHM	2.881	1.369	2.105	0.055
Spring 1st	ANYEPTOP	7.263	4.670	1.555	0.144

Table 129

One-Program Schools K-1 Analysis for Language
All Students; Program and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	233.426	11.510	20.281	0.000
Spring 1st	PGMIS	3.342	9.914	0.337	0.743
Spring 1st	PONEK1	23.887	15.218	1.570	0.148
Spring 1st	ABS1	−0.441	0.243	−1.815	0.100
Spring 1st	EDAVG	0.833	0.671	1.240	0.243
Spring 1st	PRESCHY	−2.172	4.333	−0.501	0.627
Spring 1st	FEMALE	1.691	3.692	0.458	0.657
Spring 1st	BOOKSHM	2.386	1.430	1.668	0.126
Spring 1st	ANYEPTOP	6.389	4.784	1.335	0.211

Table 130 through Table 133 show the results for the language subtest after controlling for district. The program differences now range from 7.8 to 8.1 ESS points but remain nonsignificant. The test for overall district differences is significant at the .01 level when no background variables are included (p = .003 and p = .007). When the key covariates are added or all covariates are added, district differences are reduced but are still significant at the .05 level (p = .019 and p = .039). District B has about the same level of language achievement at spring of first grade as the omitted district, district C. Districts A and F are about even with each other on the language subtest and bot' are higher than districts B and C. The composite district, H/I, has average achievement scores that lie between these two pairs of districts. H/I is somewhat closer to the higher pair (A and F) but is not quite significantly different from the reference district (p about .08 or greater).

Table 130

One-Program Schools K-1 Analysis for Language
All Students; Program and District

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	225.574	13.878	16.254	0.000
Spring 1st	PGMIS	7.772	9.643	0.806	0.435
Spring 1st	DISTA	46.309	16.226	2.854	0.014
Spring 1st	DISTB	7.460	16.501	0.452	0.659
Spring 1st	DISTF	63.593	23.664	2.687	0.019
Spring 1st	DISTHI	28.071	14.768	1.901	0.080

Table 131

One-Program Schools K-1 Analysis for Language
All Students; Program, District, and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	227.138	15.076	15.066	0.000
Spring 1st	PGMIS	8.110	9.720	0.834	0.420
Spring 1st	DISTA	46.438	16.213	2.864	0.014
Spring 1st	DISTB	6.364	16.995	0.374	0.715
Spring 1st	DISTF	66.154	25.563	2.588	0.024
Spring 1st	DISTHI	28.829	15.025	1.919	0.079
Spring 1st	PONEK1	-4.126	15.665	-0.263	0.797

ERIC

Table 132

One-Program Schools K-1 Analysis for Language
All Students; Program, District, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	227.257	14.677	15.483	0.000
Spring 1st	PGMIS	7.978	9.188	0.868	0.408
Spring 1st	DISTA	40.415	15.382	2.627	0.027
Spring 1st	DISTB	4.972	16.052	0.310	0.764
Spring 1st	DISTF	58.561	24.464	2.394	0.040
Spring 1st	DISTHI	26.217	14.255	1.839	0.099
Spring 1st	PONEK1	-3.916	15.688	-0.250	0.808
Spring 1st	ABS1	-0.407	0.242	-1.681	0.127
Spring 1st	BOOKSHM	2.559	1.371	1.867	0.095
Spring 1st	ANYEPTOP	5.923	4.686	1.264	0.238

Table 133

One-Program Schools K-1 Analysis for Language
All Students; Program, District, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	221.775	17.217	12.881	0.000
Spring 1st	PGMIS	8.067	9.130	0.884	0.411
Spring 1st	DISTA	39.936	15.302	2.610	0.040
Spring 1st	DISTB	6.912	16.471	0.420	0.689
Spring 1st	DISTF	55.380	25.054	2.210	0.069
Spring 1st	DISTHI	24.427	14.543	1.680	0.144
Spring 1st	PONEK1	2.362	20.446	0.116	0.912
Spring 1st	ABS1	-0.416	0.243	-1.711	0.138
Spring 1st	EDAVG	0.403	0.737	0.546	0.605
Spring 1st	PRESCHY	-0.743	4.438	-0.167	0.873
Spring 1st	FEMALE	1.657	3.687	0.449	0.669
Spring 1st	BOOKSHM	2.303	1.430	1.610	0.158
Spring 1st	ANYEPTOP	5.539	4.799	1.154	0.292

Summary

Although the students in immersion strategy schools tend to score about 8 points higher than students in early-exit schools on the language subtest, the difference is not statistically significant. Students in districts A and F have significantly higher first-grade language scores than students in districts B and C, with district H/I students in between.

How were the K-1 analyses for language affected by controlling for pretest?

The evaluation of the effect of including the pretest as a covariate for the language subtest parallels the evaluation for the mathematics subtest. Table 134 through Table 137 summarize the estimated models when the sample is limited to the students with a pretest and district is not controlled. Table 138 through Table 141 show the corresponding models with district controlled. These tables are based on 16 schools in district A, B, C, and H/I; as usual, district F must be omitted from pretest analyses.

Table 134

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	246.064	8.043	30.594	0.000
Spring 1st	PGMIS	9.639	13.097	0.736	0.474

Table 135

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	242.853	9.790	24.807	0.000
Spring 1st	PGMIS	8.260	13.209	0.625	0.543
Spring 1st	PONEK1	9.159	16.254	0.564	0.583

ERIC

Table 136

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	236.641	10.217	23.161	0.000
Spring 1st	PGMIS	8.404	12.444	0.675	0.515
Spring 1st	PONEK1	10.303	16.082	0.641	0.536
Spring 1st	ABS1	-0.144	0.260	-0.556	0.590
Spring 1st	BOOKSHM	2.824	1.536	1.838	0.096
Spring 1st	ANYEPTOP	6.588	5.591	1.178	0.266

Table 137

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	231.720	13.257	17.480	0.000
Spring 1st	PGMIS	8.001	12.367	0.647	0.538
Spring 1st	PONEK1	13.231	19.195	0.689	0.513
Spring 1st	ABS1	-0.143	0.261	-0.548	0.601
Spring 1st	EDAVG	0.406	0.829	0.489	0.640
Spring 1st	PRESCHY	0.357	4.968	0.072	0.945
Spring 1st	FEMALE	1.935	4.207	0.460	0.660
Spring 1st	BOOKSHM	2.525	1.632	1.547	0.166
Spring 1st	ANYEPTOP	6.001	5.771	1.040	0.333

Table 138

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program and District

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	222.737	16.070	13.860	0.000
Spring 1st	PGMIS	9.988	11.818	0.845	0.416
Spring 1st	DISTA	43.683	18.910	2.310	0.041
Spring 1st	DISTB	9.887	18.824	0.525	0.610
Spring 1st	DISTHI	27.289	17.589	1.552	0.149

Table 139

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, District, and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	222.357	17.496	12.709	0.000
Spring 1st	PGMIS	9.913	11.929	0.831	0.425
Spring 1st	DISTA	43.664	18.959	2.303	0.044
Spring 1st	DISTB	10.154	19.467	0.522	0.613
Spring 1st	DISTHI	27.129	17.867	1.518	0.160
Spring 1st	PONEK1	0.994	17.909	0.055	0.957

Table 140

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, District, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	219.617	17.528	12.530	0.000
Spring 1st	PGMIS	9.955	11.683	0.852	0.422
Spring 1st	DISTA	38.581	18.709	2.062	0.078
Spring 1st	DISTB	8.973	19.070	0.471	0.652
Spring 1st	DISTHI	23.488	17.595	1.335	0.224
Spring 1st	PONEK1	2.512	17.959	0.140	0.893
Spring 1st	ABS1	-0.131	0.260	-0.505	0.629
Spring 1st	BOOKSHM	2.538	1.543	1.645	0.144
Spring 1st	ANYEPTOP	5.529	5.619	0.984	0.358

Table 141

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, District, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	218.884	20.509	10.673	0.000
Spring 1st	PGMIS	9.944	11.644	0.854	0.441
Spring 1st	DISTA	38.397	18.691	2.054	0.109
Spring 1st	DISTB	8.509	19.595	0.434	0.686
Spring 1st	DISTHI	22.892	18.020	1.270	0.273
Spring 1st	PONEK1	1.114	23.431	0.048	0.964
Spring 1st	ABS1	-0.125	0.262	-0.477	0.659
Spring 1st	EDAVG	0.033	0.913	0.036	0.973
Spring 1st	PRESCHY	1.118	5.070	0.221	0.836
Spring 1st	FEMALE	1.951	4.208	0.464	0.667
Spring 1st	BOOKSHM	2.473	1.637	1.511	0.205
Spring 1st	ANYEPTOP	5.480	5.809	0.943	0.399

These analyses indicate an immersion strategy advantage of 8 to 10 points, but the program difference is never even close to statistical significance. This estimated program difference is comparable to the difference found in the analyses for all students when district was controlled (see Table 130 through Table 133). Although the estimated district differences are not appreciably changed by the reduction to students with pretest, they are no longer statistically significant. For the two analyses without background variables (see Table 138 and Table 139), the test for overall district differences approaches significance ($p = .061$ and $p = .077$, respectively). For the two analyses that include background variables, the district differences are reduced and are not significant ($p = .143$ and $p = .159$).

For the language subtest, the sum of the mathematics and language pretests (ANALTOBE) was found to be a better predictor of spring first grade achievement than the language pretest or the mathematics pretest alone. Table 142 through Table 145 show the results of including the pretest when district is not controlled. Table 146 through Table 149 are the corresponding tables with district included. In all of these

743

ERIC

analyses, the pretest total is included as a predictor of the propensity score (PONEK1PP). Analyses were also performed using other versions of the propensity score, but those alternative analyses were virtually identical to the analyses using PONEK1PP and are not presented here.

Table 142

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	158.518	18.608	8.519	0.000
Spring 1st	PGMIS	11.748	11.319	1.038	0.318
Spring 1st	ANALTOBE	0.249	0.049	5.066	0.000

Table 143

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, Propensity, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	142.481	21.961	6.488	0.000
Spring 1st	PGMIS	8.817	11.125	0.793	0.443
Spring 1st	PONEK1PP	21.118	16.172	1.306	0.216
Spring 1st	ANALTOBE	0.275	0.053	5.231	0.000

ERIC

Table 144

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, Key Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	138.565	22.113	6.266	0.000
Spring 1st	PGMIS	9.103	10.524	0.865	0.410
Spring 1st	PONEK1PP	20.045	15.987	1.254	0.241
Spring 1st	ABS1	-0.051	0.253	-0.201	0.846
Spring 1st	BOOKSHM	2.183	1.494	1.461	0.178
Spring 1st	ANYEPTOP	8.083	5.425	1.490	0.170
Spring 1st	ANALITOBE	0.271	0.053	5.149	0.001

Table 145

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, All Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	132.835	24.173	5.495	0.002
Spring 1st	PGMIS	8.771	10.535	0.833	0.437
Spring 1st	PONEK1PP	24.437	18.507	1.320	0.235
Spring 1st	ABS1	-0.055	0.254	-0.217	0.835
Spring 1st	EDAVG	0.377	0.774	0.487	0.644
Spring 1st	PRESCHY	-1.366	4.816	-0.284	0.786
Spring 1st	FEMALE	1.596	4.084	0.391	0.710
Spring 1st	BOOKSHM	1.874	1.592	1.177	0.284
Spring 1st	ANYEPTOP	7.505	5.614	1.337	0.230
Spring 1st	ANALITOBE	0.276	0.054	5.105	0.002

ERIC

Table 146

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, District, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	139.640	22.286	6.266	0.000
Spring 1st	PGMIS	12.105	10.546	1.148	0.278
Spring 1st	DISTA	37.502	16.816	2.230	0.050
Spring 1st	DISTB	13.197	16.710	0.790	0.448
Spring 1st	DISTHI	25.640	15.612	1.642	0.132
Spring 1st	ANALTOBE	0.240	0.049	4.852	0.001

Table 147

One-Program Schools K-1 Analysis for Language
Students with Pretest; Program, District,
Propensity, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	128.475	26.543	4.840	0.001
Spring 1st	PGMIS	11.044	10.597	1.042	0.325
Spring 1st	DISTA	36.177	16.841	2.148	0.060
Spring 1st	DISTB	16.522	17.200	0.961	0.362
Spring 1st	DISTHI	23.277	15.852	1.468	0.176
Spring 1st	PONEK1PP	14.197	18.427	0.770	0.461
Spring 1st	ANALTOBE	0.259	0.055	4.698	0.001

Table 148

One-Program Schools K-1 Analysis for Language Students with Pretest; Program, District, Key Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	126.962	26.753	4.746	0.003
Spring 1st	PGMIS	11.106	10.378	1.070	0.326
Spring 1st	DISTA	31.561	16.632	1.898	0.107
Spring 1st	DISTB	14.992	16.858	0.889	0.408
Spring 1st	DISTHI	20.284	15.608	1.300	0.241
Spring 1st	PONEK1PP	14.050	18.461	0.761	0.475
Spring 1st	ABS1	-0.040	0.253	-0.160	0.878
Spring 1st	BOOKSHM	1.968	1.501	1.312	0.238
Spring 1st	ANYEPTOP	7.074	5.465	1.294	0.243
Spring 1st	ANALTOBE	0.257	0.055	4.628	0.004

Table 149

One-Program Schools K-1 Analysis for Language Students with Pretest; Program, District, All Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	122.173	30.674	3.983	0.028
Spring 1st	PGMIS	11.071	10.387	1.066	0.365
Spring 1st	DISTA	31.197	16.748	1.863	0.159
Spring 1st	DISTB	16.210	17.367	0.933	0.419
Spring 1st	DISTHI	19.417	16.075	1.208	0.314
Spring 1st	PONEK1PP	17.983	23.301	0.772	0.496
Spring 1st	ABS1	-0.041	0.255	-0.160	0.883
Spring 1st	EDAVG	0.193	0.851	0.227	0.835
Spring 1st	PRESCHY	-1.086	4.948	-0.219	0.840
Spring 1st	FEMALE	1.656	4.090	0.405	0.713
Spring 1st	BOOKSHM	1.792	1.602	1.119	0.345
Spring 1st	ANYEPTOP	6.832	5.665	1.206	0.314
Spring 1st	ANALTOBE	0.261	0.059	4.464	0.021

Although the pretest is highly significant as a predictor of achievement, including it as a covariate changes the estimated program effect by only a point or two. The program difference remains nonsignifi-

ERIC

cant. District effects are changed somewhat, but the districts remain in the same relationship to each other. Overall district effects are not significant for any of the analyses including the pretest as a covariate (p greater than .10).

Summary

For the language subtest, reducing the sample to students with pretest scores has only small effects. The estimated program difference is about the same as for the analyses using all students with district controlled and remains nonsignificant. Differences among the four districts with pretest scores available are similar to the differences in the analyses with all students, but they are no longer statistically significant. Adding the pretest total as a covariate increases the estimated program effect by only a point or two and it remains nonsignificant. Including the pretest as a covariate has a larger impact on the estimated district effects but the districts remain in the same relationship to each other. The results of the analyses are similar with or without propensity scores and with or without background variables included as covariates.

What were the conclusions from the K-1 analyses for the language subtest?

As for the mathematics subtest, the K-1 analyses of the language subtest showed no statistically significant differences between the immersion strategy and early-exit students. However, the sign and magnitude of the program difference is consistent across most of the analyses, showing an immersion strategy advantage of about 8 to 12 points. These differences are similar in size to the differences found in the analyses for two-program schools, where the differences were also nonsignificant.

District differences were statistically significant in the K-1 analyses of all students, with districts A and F exhibiting higher language achievement levels than districts B and C. The level of language achievement in district H/I was in between and not markedly different from either pair of districts. The reduction of the sample to students with pretest scores (and therefore the elimination of district F from the comparisons) had little effect on the estimated district differences among the four remaining districts. However, the district differences no longer achieved statistical significance at the .05 level. The district differences were reduced by the inclusion of additional covariates (either background variables or the pretest), but the relationships among the districts and the statistical significance of the district differences were not appreciably affected.

It is interesting to relate the district differences on the language subtest to the patterns of instructional English use described in Chapter I. District F, with the greatest use of English in its early-exit program so that it more closely resembled an immersion strategy program, had the highest estimated level of first-grade language achievement scores. With only 24 students included in these K-1 analyses, however, that estimate is highly variable. Districts A and H/I, with the next greatest use of English in their early-exit schools, had the next highest levels of language achievement in first grade. Districts B and C, with lower use of English in their early-exit schools, had the lowest estimated levels of language achievement in first grade. The association between the amount of English used for instruction and the level of achievement on the language subtest in first grade is intriguing, but a formal evaluation of this possible relationship goes beyond the scope of nominal program analyses.

Results of the K-1 Analyses for Reading

What were the results of the basic K-1 analyses for the reading subtest?

Table 150 through Table 153 give the results of the reading subtest analyses without controlling for district, and Table 154 through Table 157 present the district-controlled results. In all eight analyses the immersion strategy students have slightly lower scores than early-exit students, with the difference ranging from less than 2 points to about 3.6 points. These small differences do not approach statistical significance.

Table 150

One-Program Schools K-1 Analysis for Reading
All Students; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	237.766	5.062	46.969	0.000
Spring 1st	PGMIS	-3.598	8.294	-0.434	0.670

Table 151

One-Program Schools K-1 Analysis for Reading
All Students; Program and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	238.466	6.970	34.215	0.000
Spring 1st	PGMIS	-3.425	8.504	-0.403	0.693
Spring 1st	PONEK1	-1.642	10.975	-0.150	0.883

Table 152

One-Program Schools K-1 Analysis for Reading
All Students; Program and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	[illegible]SE	237.045	7.310	32.427	0.000
Spring 1st	PGMI[illegible]	-2.586	7.856	-0.329	0.747
Spring 1st	PONEK1	-3.054	10.663	-0.286	0.779
Spring 1st	ABS1	-0.476	0.205	-2.325	0.037
Spring 1st	BOOKSHM	2.025	1.158	1.749	0.104
Spring 1st	ANYEPTOP	8.392	3.954	2.123	0.054

Table 153

One-Program Schools K-1 Analysis for Reading
All Students; Program and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	228.822	9.167	24.961	0.000
Spring 1st	PGMIS	-2.656	7.512	-0.354	0.731
Spring 1st	PONEK1	-3.119	12.242	-0.255	0.804
Spring 1st	ABS1	-0.488	0.205	-2.382	0.039
Spring 1st	EDAVG	0.801	0.559	1.434	0.182
Spring 1st	PRESCHY	5.421	3.651	1.485	0.168
Spring 1st	FEMALE	1.993	3.122	0.638	0.538
Spring 1st	BOOKSHM	1.560	1.207	1.292	0.225
Spring 1st	ANYEPTOP	7.119	4.035	1.764	0.108

Table 154

One-Program Schools K-1 Analysis for Reading
All Students; Program and District

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	224.584	12.548	17.898	0.000
Spring 1st	PGMIS	-2.849	8.685	-0.328	0.748
Spring 1st	DISTA	24.812	14.668	1.692	0.115
Spring 1st	DISTB	3.651	14.890	0.245	0.810
Spring 1st	DISTF	18.458	21.342	0.865	0.403
Spring 1st	DISTHI	14.059	13.339	1.054	0.311

Table 155

One-Program Schools K-1 Analysis for Reading
All Students; Program, District, and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	229.191	13.204	17.357	0.000
Spring 1st	PGMIS	-1.869	8.530	-0.219	0.830
Spring 1st	DISTA	25.187	14.264	1.766	0.103
Spring 1st	DISTB	0.404	14.912	0.027	0.979
Spring 1st	DISTF	25.970	22.370	1.161	0.268
Spring 1st	DISTHI	16.320	13.200	1.236	0.240
Spring 1st	PONEK1	-12.119	13.375	-0.906	0.383

Table 156

One-Program Schools K-1 Analysis for Reading
All Students; Program, District, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	231.206	12.990	17.799	0.000
Spring 1st	PGMIS	-2.019	8.177	-0.247	0.811
Spring 1st	DISTA	19.447	13.739	1.416	0.191
Spring 1st	DISTB	-1.020	14.283	-0.071	0.945
Spring 1st	DISTF	18.411	21.652	0.850	0.417
Spring 1st	DISTHI	14.387	12.702	1.133	0.287
Spring 1st	PONEK1	-12.844	13.354	-0.962	0.361
Spring 1st	ABS1	-0.477	0.206	-2.320	0.046
Spring 1st	BOOKSHM	1.774	1.166	1.521	0.163
Spring 1st	ANYEPTOP	7.663	3.986	1.923	0.087

Table 157

One-Program Schools K-1 Analysis for Reading
All Students; Program, District, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	228.041	14.985	15.218	0.000
Spring 1st	PGMIS	-1.986	8.071	-0.246	0.814
Spring 1st	DISTA	18.269	13.569	1.346	0.227
Spring 1st	DISTB	-2.763	14.518	-0.190	0.855
Spring 1st	DISTF	16.488	21.982	0.750	0.482
Spring 1st	DISTHI	11.855	12.840	0.923	0.391
Spring 1st	PONEK1	-16.483	17.324	-0.951	0.378
Spring 1st	ABS1	-0.474	0.206	-2.301	0.061
Spring 1st	EDAVG	0.470	0.625	0.751	0.481
Spring 1st	PRESCHY	6.441	3.760	1.713	0.138
Spring 1st	FEMALE	1.917	3.124	0.614	0.562
Spring 1st	BOOKSHM	1.496	1.212	1.234	0.263
Spring 1st	ANYEPTOP	6.772	4.067	1.665	0.147

The "key covariates" for the reading subtest are the same as for the language subtest: ABS1, BOOKSHM, and ANYEPTOP. For the language subtest, BOOKSHM was found to be slightly more powerful as a predictor, but for the reading subtest ANYEPTOP was the more powerful. To facilitate comparisons, both variables were included as key covariates for both subtests. As for the mathematics and language subtests, the estimated program difference is little affected by the inclusion of propensity score and background variables.

The pattern of district differences for the reading subtest is similar to that found for the language subtest: districts A and F have the highest estimates, districts B and C the lowest estimates, and district H/I is in between (but closer to districts A and F). However, the district differences for the reading subtest are never statistically significant (p greater than .10). That is, the hypothesis of no district differences on the reading subtest cannot be rejected.

Summary

The immersion strategy students tend to score a few points lower than early-exit students on the reading subtest but the difference does not approach statistical significance. Although the pattern of district achievement levels on the reading subtest in first grade is similar to that found for the language subtest, the district differences are not statistically significant.

How were the K-1 analyses for reading affected by controlling for pretest?

For the reading subtest, reducing the sample to the students with pretest scores has the effect of changing the estimated program effect to a small immersion strategy advantage of 3.8 to 4.9 points (see Table 158 through Table 165). This small difference is almost unaffected by controlling for district and never approaches statistical significance. The estimated district effects among the four remaining districts have the same pattern as in the analyses of all students and remain nonsignificant.

Table 158

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	232.764	6.561	35.476	0.000
Spring 1st	PGMIS	4.615	10.684	0.432	0.672

Table 159

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	231.616	8.226	28.158	0.000
Spring 1st	PGMIS	4.138	10.967	0.377	0.712
Spring 1st	PONEK1	3.268	14.019	0.233	0.819

Table 160

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	230.671	8.619	26.763	0.000
Spring 1st	PGMIS	4.457	10.335	0.431	0.675
Spring 1st	PONEK1	2.975	13.817	0.215	0.834
Spring 1st	ABS1	-0.398	0.225	-1.769	0.107
Spring 1st	BOOKSHM	1.320	1.331	0.991	0.345
Spring 1st	ANYEPTOP	9.111	4.846	1.880	0.090

Table 161

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	217.818	11.147	19.541	0.000
Spring 1st	PGMIS	3.774	10.247	0.368	0.724
Spring 1st	PONEK1	1.465	16.253	0.090	0.931
Spring 1st	ABS1	-0.373	0.224	-1.670	0.139
Spring 1st	EDAVG	1.070	0.706	1.514	0.174
Spring 1st	PRESCHY	7.358	4.248	1.732	0.127
Spring 1st	FEMALE	6.766	3.601	1.879	0.102
Spring 1st	BOOKSHM	0.513	1.397	0.368	0.724
Spring 1st	ANYEPTOP	7.206	4.939	1.459	0.188

Table 162

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and District

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	220.749	16.096	13.715	0.000
Spring 1st	PGMIS	4.732	11.733	0.403	0.694
Spring 1st	DISTA	19.464	18.890	1.030	0.325
Spring 1st	DISTB	6.297	18.807	0.335	0.744
Spring 1st	DISTHI	15.306	17.560	0.872	0.402

Table 163

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, District, and Propensity

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	221.344	17.162	12.897	0.000
Spring 1st	PGMIS	4.851	11.800	0.411	0.690
Spring 1st	DISTA	19.496	18.895	1.032	0.326
Spring 1st	DISTB	5.879	19.268	0.305	0.767
Spring 1st	DISTHI	15.557	17.744	0.877	0.401
Spring 1st	PONEK1	-1.556	15.623	-0.100	0.923

Table 164

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, District, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	223.174	16.885	13.217	0.000
Spring 1st	PGMIS	4.885	11.392	0.429	0.681
Spring 1st	DISTA	14.127	18.343	0.770	0.466
Spring 1st	DISTB	4.688	18.605	0.252	0.808
Spring 1st	DISTHI	13.554	17.200	0.788	0.457
Spring 1st	PONEK1	-1.769	15.604	-0.113	0.913
Spring 1st	ABS1	-0.396	0.226	-1.752	0.123
Spring 1st	BOOKSHM	1.171	1.341	0.873	0.412
Spring 1st	ANYEPTOP	8.615	4.878	1.766	0.121

Table 165

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, District, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	212.632	19.189	11.081	0.000
Spring 1st	PGMIS	4.861	11.372	0.427	0.691
Spring 1st	DISTA	11.570	18.355	0.630	0.563
Spring 1st	DISTB	4.355	19.018	0.229	0.830
Spring 1st	DISTHI	6.760	17.537	0.385	0.719
Spring 1st	PONEK1	-1.058	20.097	-0.053	0.961
Spring 1st	ABS1	-0.364	0.224	-1.620	0.181
Spring 1st	EDAVG	0.988	0.782	1.263	0.275
Spring 1st	PRESCHY	7.494	4.346	1.724	0.160
Spring 1st	FEMALE	6.776	3.608	1.878	0.134
Spring 1st	BOOKSHM	0.487	1.405	0.346	0.746
Spring 1st	ANYEPTOP	6.873	4.979	1.380	0.240

As for the language pretest, the pretest total (ANALTOBE) was found to be the pretest score with greatest predictive power. Table 166 through Table 173 show the effect of including the pretest total as a predictor of first grade reading achievement and as a predictor of propensity. The estimated program difference is increased somewhat by the inclusion of the pretest total as a predictor, to a range of 4.9 to 6.2 points, but the difference still falls well short of statistical significance.

Table 166

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	172.720	16.308	10.591	0.000
Spring 1st	PGMIS	5.849	9.663	0.605	0.555
Spring 1st	ANALTOBE	0.171	0.043	3.959	0.002

ERIC

Table 167

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, Propensity, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	167.892	19.380	8.663	0.000
Spring 1st	PGMIS	4.936	9.849	0.501	0.625
Spring 1st	PONEK1PP	6.564	14.272	0.460	0.654
Spring 1st	ANALTOBE	0.179	0.046	3.858	0.002

Table 168

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, Key Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	169.841	19.478	8.719	0.000
Spring 1st	PGMIS	5.390	9.365	0.575	0.579
Spring 1st	PONEK1PP	4.456	14.087	0.316	0.759
Spring 1st	ABS1	-0.337	0.222	-1.518	0.163
Spring 1st	BOOKSHM	0.905	1.313	0.690	0.508
Spring 1st	ANYEPTOP	9.977	4.765	2.094	0.066
Spring 1st	ANALTOBE	0.173	0.046	3.729	0.005

Table 169

e-Program Schools K-1 Analysis for Reading
Students w .n Pretest; Program, All Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	162.868	21.124	7.710	0.000
Spring 1st	PGMIS	4.969	9.341	0.532	0.614
Spring 1st	PONEK1PP	0.636	16.188	0.039	0.970
Spring 1st	ABS1	-0.317	0.221	-1.437	0.201
Spring 1st	EDAVG	0.935	0.674	1.387	0.215
Spring 1st	PRESCHY	6.962	4.184	1.664	0.147
Spring 1st	FEMALE	6.598	3.545	1.861	0.112
Spring 1st	BOOKSHM	0.206	1.382	0.149	0.887
Spring 1st	ANYEPTOP	8.295	4.875	1.702	0.140
Spring 1st	ANALTOBE	0.161	0.047	3.435	0.014

ERIC

Table 170

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, District, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	163.283	21.185	7.708	0.000
Spring 1st	PGMIS	6.170	10.846	0.569	0.582
Spring 1st	DISTA	15.361	17.436	0.881	0.399
Spring 1st	DISTB	8.585	17.333	0.495	0.631
Spring 1st	DISTHI	14.179	16.181	0.876	0.401
Spring 1st	ANALTOBE	0.166	0.044	3.801	0.003

Table 171

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, District,
Propensity, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	160.780	24.734	6.500	0.000
Spring 1st	PGMIS	5.926	10.905	0.543	0.600
Spring 1st	DISTA	15.073	17.484	0.862	0.411
Spring 1st	DISTB	9.330	17.735	0.526	0.612
Spring 1st	DISTHI	13.646	16.392	0.832	0.427
Spring 1st	PONEK1PP	3.184	16.279	0.196	0.849
Spring 1st	ANALTOBE	0.170	0.049	3.499	0.007

Table 172

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, District,
Key Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	164.510	24.675	6.667	0.001
Spring 1st	PGMIS	5.989	10.536	0.568	0.590
Spring 1st	DISTA	10.152	16.9[illegible]3	0.597	0.572
Spring 1st	DISTB	7.849	17.144	0.458	0.663
Spring 1st	DISTHI	12.144	15.896	0.764	0.474
Spring 1st	PONEK1PP	1.505	16.242	0.093	0.929
Spring 1st	ABS1	-0.337	0.223	-1.512	0.181
Spring 1st	BOOKSHM	0.799	1.321	0.605	0.567
Spring 1st	ANYEPTOP	9.659	4.802	2.011	0.091
Spring 1st	ANALTOBE	0.166	0.049	3.399	0.015

Table 173

One-Program Schools K-1 Analysis for Reading
Students with Pretest; Program, District,
All Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	157.608	27.762	5.677	0.011
Spring 1st	PGMIS	6.055	10.559	0.573	0.607
Spring 1st	DISTA	8.055	17.113	0.471	0.670
Spring 1st	DISTB	7.098	17.551	0.404	0.713
Spring 1st	DISTHI	6.120	16.269	0.376	0.732
Spring 1st	PONEK1PP	0.591	20.253	0.029	0.979
Spring 1st	ABS1	-0.310	0.222	-1.402	0.256
Spring 1st	EDAVG	0.924	0.739	1.250	0.300
Spring 1st	PRESCHY	6.863	4.297	1.597	0.209
Spring 1st	FEMALE	6.604	3.553	1.859	0.160
Spring 1st	BOOKSHM	0.162	1.394	0.116	0.915
Spring 1st	ANYEPTOP	7.988	4.920	1.624	0.203
Spring 1st	ANALTOBE	0.159	0.051	3.113	0.053

Summary

Limiting the reading subtest analyses to students with pretest scores reverses the estimated program effect from favoring early-exit to favoring immersion strategy (compare PGMIS in Table 150 through Table 157 with PGMIS in Table 158 through Table 165), but the estimated differences are no more than six ESS points and never approach statistical significance. Adding the pretest as a predictor has little additional effect. District differences are somewhat reduced and remain nonsignificant.

What were the conclusions from the K-1 analyses for the reading subtest?

For the reading subtest, there are no statistically significant differences between the programs in first grade. The estimated difference among all students shows an early-exit advantage of no more than 3.6 points; for students with pretest scores, and therefore omitting district F, the estimated difference favors immersion strategy by no more than 6.2 points. These small differences are not even close to statistical significance.

District differences for the reading subtest are smaller than for the language subtest, although they have the same pattern, and are not statistically significant.

Results of the Basic 1-3 Analyses

What were the results of the basic 1-3 analysis for the mathematics subtest?

For the 1-3 analyses of one-program schools, as for the corresponding analyses of two-program schools, first through third grade information was used to estimate individual growth curves, using the HLM computer program. (Please note that when schools are treated as a random effect in the 1-3 HLM analyses, this is a three-level model wherein growth curves are generated for individual students, schools, and overall. When schools are

ERIC

entered as a fixed effect in the 1-3 HLM analyses, this is a two-level model wherein individual student and overall growth curves are generated.) Each student with a first grade test score and a second grade or third grade test score had an individual growth curve estimated. An overall coefficient of curvature was estimated, and an intercept (status at spring of first grade) and slope (growth rate from spring of first grade through spring of third grade) was estimated for each individual. These individual intercepts and slopes were then predicted using the propensity score, background variables, a set of variables ("dummy variables") for indicating the school, and a single binary variable for the difference between the immersion strategy and early-exit programs. The school and program variables are included as predictors of both the initial status and the growth rate.

Table 174 summarizes the results for the mathematics subtest. The table may be interpreted the same way as the tables of HLM results for the 1-3 analyses given in Chapter III. The predictors include the BASE (overall constant), PGMIS (a binary variable that is 1 for IS students and 0 for EE students), EDAVG (parents' educational average), and FEMALE (a binary variable that is 1 for girls and 0 for boys). In addition, PONE13 (the propensity score) and a series of variables beginning NIS or NEE (school variables, some of which are groups of schools; see page 173 for discussion of school groups) are used to predict both slope and intercept.

Table 174

One-Program Schools 1-3 Analysis:
Basic Two-Level Model for Mathematics

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	256.447	7.849	32.672	0.000
Spring 1st	PGMIS	3.179	3.184	0.999	0.319
Spring 1st	PONE13	-14.353	11.453	-1.253	0.211
Spring 1st	NIS01	-9.108	3.753	-2.427	0.016
Spring 1st	NIS10	-5.119	9.163	-0.559	0.577
Spring 1st	NIS7A	17.363	5.039	3.446	0.001
Spring 1st	NIS7B	-2.825	4.731	-0.597	0.551
Spring 1st	NIS7C	2.198	7.381	0.298	0.766
Spring 1st	NIS7D	-8.585	6.030	-1.424	0.155
Spring 1st	NEE02	17.397	4.743	3.668	0.000
Spring 1st	NEE04	0.618	7.518	0.082	0.934
Spring 1st	NEE05	3.424	4.851	0.706	0.481
Spring 1st	NEE11	-12.150	6.682	-1.818	0.070
Spring 1st	NEE13	-18.983	7.062	-2.688	0.007
Spring 1st	NEE15	-16.848	7.163	-2.352	0.019
Spring 1st	NEE50	19.580	8.226	2.380	0.018
Spring 1st	NEE81	-2.502	7.219	-0.347	0.729
Spring 1st	NEE82	39.093	6.990	5.593	0.000
Spring 1st	NEE83	-10.931	5.110	-2.139	0.033
Spring 1st	NEE84	-13.518	8.595	-1.573	0.117
Spring 1st	EDAVG	1.102	0.444	2.482	0.013
Spring 1st	FEMALE	-4.781	2.576	-1.856	0.064
Growth Rate 1-3	BASE	47.123	3.357	14.036	0.000
Growth Rate 1-3	PGMIS	-3.183	2.186	-1.456	0.146
Growth Rate 1-3	PONE13	-4.177	6.970	-0.599	0.549
Growth Rate 1-3	NIS01	5.122	2.409	2.126	0.034
Growth Rate 1-3	NIS10	-2.507	5.866	-0.427	0.669
Growth Rate 1-3	NIS7A	9.337	3.492	2.674	0.008
Growth Rate 1-3	NIS7B	2.477	3.401	0.728	0.467
Growth Rate 1-3	NIS7C	-2.105	4.535	-0.464	0.643
Growth Rate 1-3	NIS7D	1.719	4.359	0.394	0.693
Growth Rate 1-3	NEE02	-0.012	3.106	-0.004	0.997
Growth Rate 1-3	NEE04	14.855	4.598	3.231	0.001
Growth Rate 1-3	NEE05	4.783	3.042	1.573	0.117
Growth Rate 1-3	NEE11	-16.980	4.069	-4.173	0.000
Growth Rate 1-3	NEE13	11.700	4.535	2.580	0.010
Growth Rate 1-3	NEE15	-18.850	4.228	-4.458	0.000
Growth Rate 1-3	NEE50	-1.940	5.107	-0.380	0.704
Growth Rate 1-3	NEE81	-1.907	6.031	-0.316	0.752
Growth Rate 1-3	NEE82	-14.004	5.022	-2.789	0.006
Growth Rate 1-3	NEE83	19.101	3.848	4.964	0.000
Growth Rate 1-3	NEE84	11.070	5.197	2.130	0.034
Growth Rate 1-3	FEMALE	5.713	1.739	3.284	0.001
Curvature 1-3	BASE	-6.197	2.708	-2.289	0.023
Curvature 1-3	PGMIS	-3.657	2.624	-1.394	0.164
Curvature 1-3	PONE13	11.687	5.071	2.305	0.022
Year-To-Year Deviations	ABSENT*	-0.342	0.123	-2.769	0.006

* Variable represents days absent from school.

263

The NIS and NEE variables are indicators of the child's school or school group. The district C schools (the group called NIS2A or NEE2A) were the omitted school groups in this parametrization. Thus a child in an immersion strategy school in district C would have the value -1 on all the NIS school variables, and a child in an early-exit school in district C would have the value -1 on all the NEE school variables. The parameters for the NIS variables represent the difference between that school group and the average of all the immersion strategy school groups. The difference between the omitted school group, NIS2A, and the average of all the IS schools can be calculated as the negative of the sum of the coefficients for the NIS variables. Any IS school group could have been selected as the omitted school group; the choice does not affect the statistical results. Similarly, the parameters for the NEE variables represent the difference between that school group and the average of all the early-exit school groups. The omitted school group is NEE2A, the early-exit schools in district C.

As for the two-program schools analyses, the parameter value for PGMIS indicates the increase (or decrease) in the predicted value for a student in an immersion strategy program. The value 3.179 for the parameter PGMIS for predicting initial status means the initial status of immersion strategy students is predicted to be about 3.2 ESS points above that of early-exit students. This estimate reflects the adjustment for propensity. Not only is this difference not statistically significant, it is much smaller than many of the estimated differences among the school groups.

The t-statistic and associated p-value for each NIS variable tests the hypothesis that the school indicated has a parameter value of zero. This is equivalent to the hypothesis that the students in this immersion strategy school group have the same average initial status as the average of all the immersion strategy school groups. The t-statistics for the NEE variables are interpreted similarly. The omitted school groups (NIS2A and NEE2A) can be tested using multivariate hypothesis tests. Neither is

ERIC

significant for predicting initial status (p = .269 and p = .281, respectively).

For the growth rate, the programs are not significantly different (difference = -3.183, p = .146) after adjusting for the propensity score (PONE13). The schools and school groups differ considerably in the average growth rate of their students, however. As mentioned above, these school differences may be attributable to differences in program implementation, but analyses to relate the school differences to information about the schools go beyond the scope of the nominal program analyses of this report. The NIS2A school group has significantly lower growth than the average IS school (difference = -14.043, p less than .0005). The NEE2A schools also exhibit lower growth than the average EE school (difference = -7.816, p = .025).

Unlike the students in the two-program schools, the students in these one-program schools exhibit significant variation in curvature on the mathematics subtest. After adjusting for the propensity score, the program difference in curvature is not significant (difference = -3.657, p = .164). A simultaneous test of the two predictors of curvature, PGMIS and PONE13, just reaches statistical significance at the .05 level (p = .047). Together with the fact that the effect of propensity score alone is not highly significant (p = .022), this means that the amount of student variation in curvature is not large. To the extent that there is any difference between the programs, the immersion strategy students exhibit somewhat greater negative curvature (i.e., more flattening). It should be emphasized that this difference is not statistically significant after adjusting for the propensity score.

The coefficient for the number of absences for predicting year-to-year deviations in the test scores is -0.342, which is statistically significant (p = .006). This means that for every day absent in a year, mathematics test scores are predicted to be .342 ESS points lower. As discussed in Chapter II, using absences to predict year-to-year deviations is an alternative to including the average number of absences as a predictor of initial status (spring of first grade).

Summary

Consistent with the hypothesized growth for each program and with the two-program school analyses, this basic two-level model uncovers no statistically significant differences between the programs on the mathematics subtest for the one-program schools after adjusting for propensity. The schools within each program differ significantly from each other, even within the same district.

What were the results of the basic 1-3 analysis for the language subtest?

Table 175 presents the results of the basic 1-3 analysis for the language subtest. The model is the same as for the mathematics subtest. Besides the propensity score (PONE13), only parents' education (EDAVG) and students' gender (FEMALE) are used as covariates. Notice that although girls do not differ significantly from boys in initial status (difference = 3.202, p = .431), girls have a substantially higher growth rate (difference = 14.581, p less than .0005). Unlike mathematics, the number of absences does not have a significant effect on test scores for language (p = .497). This is consistent with the expectation that students learn language in school and out of school but tend to learn mathematics mostly in school.

Table 175

One-Program Schools 1-3 Analysis:
Basic Two-Level Model for Language

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	230.929	12.586	18.348	0.000
Spring 1st	PGMIS	3.817	5.023	0.760	0.448
Spring 1st	PONE13	26.862	18.174	1.478	0.140
Spring 1st	NIS01	2.243	5.916	0.379	0.705
Spring 1st	NIS10	16.816	14.472	1.162	0.246
Spring 1st	NIS7A	26.072	7.983	3.266	0.001
Spring 1st	NIS7B	-8.060	7.497	-1.075	0.283
Spring 1st	NIS7C	-13.108	11.641	-1.126	0.261
Spring 1st	NIS7D	0.388	9.536	0.041	0.968
Spring 1st	NEE02	45.790	7.474	6.126	0.000
Spring 1st	NEE04	28.460	11.841	2.404	0.017
Spring 1st	NEE05	16.101	7.645	2.106	0.036
Spring 1st	NEE11	-12.859	10.564	-1.217	0.224
Spring 1st	NEE13	-8.932	11.139	-0.802	0.423
Spring 1st	NEE15	-14.188	11.333	-1.252	0.211
Spring 1st	NEE50	12.354	12.999	0.950	0.342
Spring 1st	NEE81	-11.956	11.393	-1.049	0.295
Spring 1st	NEE82	20.078	11.033	1.820	0.070
Spring 1st	NEE83	-10.563	8.074	-1.308	0.192
Spring 1st	NEE84	-37.953	13.550	-2.801	0.005
Spring 1st	EDAVG	1.765	0.719	2.454	0.015
Spring 1st	FEMALE	3.202	4.062	0.788	0.431
Growth Rate 1-3	BASE	68.256	5.908	11.553	0.000
Growth Rate 1-3	PGMIS	4.710	3.858	1.221	0.223
Growth Rate 1-3	PONE13	-27.483	12.276	-2.239	0.026
Growth Rate 1-3	NIS01	1.541	4.269	0.361	0.718
Growth Rate 1-3	NIS10	-15.107	10.405	-1.452	0.147
Growth Rate 1-3	NIS7A	23.652	6.112	3.870	0.000
Growth Rate 1-3	NIS7B	19.863	5.921	3.355	0.001
Growth Rate 1-3	NIS7C	-12.319	8.055	-1.529	0.127
Growth Rate 1-3	NIS7D	-14.134	7.616	-1.856	0.064
Growth Rate 1-3	NEE02	-6.752	5.489	-1.230	0.219
Growth Rate 1-3	NEE04	9.053	8.176	1.107	0.269
Growth Rate 1-3	NEE05	5.818	5.391	1.079	0.281
Growth Rate 1-3	NEE11	-20.379	7.215	-2.824	0.005
Growth Rate 1-3	NEE13	-3.318	8.032	-0.413	0.680
Growth Rate 1-3	NEE15	-22.830	7.509	-3.041	0.003
Growth Rate 1-3	NEE50	8.773	9.032	0.971	0.332
Growth Rate 1-3	NEE81	5.652	10.369	0.545	0.586
Growth Rate 1-3	NEE82	12.151	8.752	1.388	0.166
Growth Rate 1-3	NEE83	13.288	6.667	1.993	0.047
Growth Rate 1-3	NEE84	17.751	9.201	1.929	0.054
Growth Rate 1-3	FEMALE	14.581	3.069	4.751	0.000
Curvature 1-3	BASE	-14.803	4.393	-3.370	0.001
Curvature 1-3	PGMIS	-14.086	4.276	-3.294	0.001
Curvature 1-3	PONE13	24.815	8.216	3.020	0.003
Year-To-Year Deviations	ABSENT	-0.137	0.202	-0.679	0.497

As expected, the school groups show considerable variation in initial status and growth rate. The schools in district C have lower initial status than average: for NIS2A, the difference is -24.351 (p = .005) and for NEE2A the difference is -26.332 (p = .001). The students in the immersion strategy schools in district C have average growth rates (difference = -3.496, p = .607) but the students in the early-exit schools in district C have below-average growth rates (difference = -26.332, p = .002) among early-exit schools.

After adjusting for propensity, there are no significant differences between the programs in initial status (difference = 3.817, p = .448) or in growth rate (difference = 4.710, p = .223). However, even after accounting for propensity score, there is a significant difference in curvature between the programs (difference = -14.086, p = .001). The difference is in the direction of more negative curvature for the immersion strategy program. Taken together with the absence of significant differences in initial status or growth rate, this implies a temporary increase in test scores for immersion strategy students, just as was found for language and reading among the students in two-program schools. (The temporary increase for language was at first grade and the temporary increase for reading was at second grade for the two-program schools.) The estimated program difference at spring of first grade is -0.878 (p = .865), at second grade it is 17.918 (p = .003), and at third grade it is 8.542 (p = .307).

Summary

The one-program schools exhibit considerable variation on the language subtest. After adjusting for propensity, as hypothesized no significant program differences in initial status or average growth were detected. However, consistent with the two-program analyses, the IS program had greater curvature than the early-exit program for language, even after adjusting for propensity. This difference in curvature indicates that at third grade the immersion strategy students in the one-program schools have slightly higher language subtest scores but flatter growth than the students in early-exit schools. While it would be tempting to extrapolate this result beyond third grade, the interpretation of these growth curves must be limited to grades one through three.

What were the results of the basic 1-3 analysis for the reading subtest?

The results of the basic 1-3 analysis for the reading subtest are summarized in Table 176. The model is the same as for the mathematics and language subtests. For the reading subtest, unlike the language subtest, the growth rate for girls is not significantly greater than for boys (difference = 4.552, p = .081). Like the language subtest, the reading subtest shows significant school variability and no significant program differences in initial status (difference = -5.896, p = .163) and average growth rate (difference = 3.251, p = .321) after adjusting for propensity.

Although the IS students show greater negative curvature for reading after adjusting for propensity (difference = -4.118), unlike for language the difference is not significant (p = .283). The estimated program differences at second grade (difference = 0.100, p = .984) and third grade (difference = -0.767, p = .914) are very small. However, the estimated program difference at spring of first grade (-7.269) begins to approach statistical significance (p = .095). This difference indicates the possibility of a small advantage for early-exit students, the same finding as for the reading subtest for the two-program schools.

Table 176

One-Program Schools 1-3 Analysis:
Basic Two-Level Model for Reading

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	225.596	10.604	21.275	0.000
Spring 1st	PGMIS	-5.896	4.222	-1.396	0.163
Spring 1st	PONE13	6.889	15.287	0.451	0.653
Spring 1st	NIS01	1.425	4.969	0.287	0.774
Spring 1st	NIS10	20.173	12.141	1.662	0.097
Spring 1st	NIS7A	28.755	6.730	4.272	0.000
Spring 1st	NIS7B	-0.471	6.323	-0.074	0.941
Spring 1st	NIS7C	-28.362	9.795	-2.896	0.004
Spring 1st	NIS7D	-1.358	8.037	-0.169	0.866
Spring 1st	NEE02	33.655	6.277	5.361	0.000
Spring 1st	NEE04	6.791	9.936	0.683	0.495
Spring 1st	NEE05	10.743	6.421	1.673	0.095
Spring 1st	NEE11	-9.856	8.865	-1.112	0.267
Spring 1st	NEE13	-0.276	9.339	-0.030	0.976
Spring 1st	NEE15	-6.727	9.513	-0.707	0.480
Spring 1st	NEE50	-1.927	10.925	-0.176	0.860
Spring 1st	NEE81	-4.816	9.577	-0.503	0.615
Spring 1st	NEE82	14.595	9.278	1.573	0.116
Spring 1st	NEE83	-7.125	6.785	-1.050	0.294
Spring 1st	NEE84	-23.607	11.395	-2.072	0.039
Spring 1st	EDAVG	1.507	0.607	2.483	0.013
Spring 1st	FEMALE	3.817	3.413	1.118	0.264
Growth Rate 1-3	BASE	60.460	5.026	12.029	0.000
Growth Rate 1-3	PGMIS	3.251	3.277	0.992	0.322
Growth Rate 1-3	PONE13	-13.534	10.436	-1.297	0.195
Growth Rate 1-3	NIS01	0.099	3.614	0.027	0.978
Growth Rate 1-3	NIS10	-7.068	8.804	-0.803	0.423
Growth Rate 1-3	NIS7A	13.496	5.219	2.586	0.010
Growth Rate 1-3	NIS7B	4.127	5.078	0.813	0.417
Growth Rate 1-3	NIS7C	-2.392	6.807	-0.351	0.725
Growth Rate 1-3	NIS7D	-19.325	6.513	-2.967	0.003
Growth Rate 1-3	NEE02	-6.011	4.655	-1.291	0.197
Growth Rate 1-3	NEE04	11.937	6.903	1.729	0.085
Growth Rate 1-3	NEE05	0.720	4.562	0.158	0.875
Growth Rate 1-3	NEE11	-15.830	6.107	-2.592	0.010
Growth Rate 1-3	NEE13	4.254	6.797	0.626	0.532
Growth Rate 1-3	NEE15	-20.621	6.348	-3.249	0.001
Growth Rate 1-3	NEE50	7.225	7.655	0.944	0.346
Growth Rate 1-3	NEE81	10.546	8.969	1.176	0.240
Growth Rate 1-3	NEE82	7.612	7.505	1.014	0.311
Growth Rate 1-3	NEE83	10.603	5.738	1.848	0.065
Growth Rate 1-3	NEE84	5.940	7.789	0.763	0.446
Growth Rate 1-3	FEMALE	4.552	2.606	1.747	0.081
Curvature 1-3	BASE	-17.453	3.944	-4.425	0.000
Curvature 1-3	PGMIS	-4.118	3.829	-1.076	0.283
Curvature 1-3	PONE13	23.445	7.383	3.176	0.002
Year-To-Year Deviations	ABSENT	-0.107	0.174	-0.614	0.539

ERIC

Recall from Chapter III, however, that for the two-program schools, there was significantly different curvature between the two programs for the reading subtest, resulting in the finding that the IS students had higher reading achievement at second grade but lower growth from second to third grade, suggesting a transitory boost in second grade and deceleration in growth in reading in third grade. For the one-program schools in these analyses, the IS and EE programs show virtually identical achievement levels in second and third grade, with early-exit slightly ahead (but not significantly ahead) at spring of first grade. Failure to find comparable results in reading growth between the one-program school and the two-program school analyses could partly be the result of not being able to control for school in the one-program school analyses.

Summary

The one-program schools exhibit considerable variation on the reading subtest but the program differences are small after adjusting for propensity. There is some evidence that the early-exit students are slightly ahead in reading at spring of first grade, but the difference is not statistically significant. The two programs show no differences in reading achievement at second and third grade for these one-program schools.

Allowing Curvature to Vary by School in the 1-3 Analyses

Does the curvature vary among the one-program schools?

The basic 1-3 analyses of one-program schools discussed earlier does not provide for the curvature to vary from school to school. This subsequent analysis presents the two-level HLM program allowing school to vary. The decision not to include school as a predictor of curvature was reached only after evaluating the amount of school variability in curvature. Table 177, Table 178, and Table 179 present the two-level HLM model estimated with curvature allowed to vary by school. The other components of the model are unchanged.

Table 177

One-Program Schools 1-3 Analysis:
School Curvature Model for Mathematics

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	256.208	7.878	32.521	0.000
Spring 1st	PGMIS	3.587	3.204	1.120	0.264
Spring 1st	PONE13	-14.658	11.507	-1.274	0.203
Spring 1st	NIS01	-9.000	3.774	-2.385	0.018
Spring 1st	NIS10	-5.308	9.203	-0.577	0.564
Spring 1st	NIS7A	17.276	5.055	3.418	0.001
Spring 1st	NIS7B	-2.905	4.762	-0.610	0.542
Spring 1st	NIS7C	2.457	7.415	0.331	0.741
Spring 1st	NIS7D	-7.953	6.064	-1.311	0.190
Spring 1st	NEE02	17.779	4.771	3.726	0.000
Spring 1st	NEE04	1.534	7.560	0.203	0.839
Spring 1st	NEE05	3.767	4.877	0.772	0.440
Spring 1st	NEE11	-11.939	6.715	-1.778	0.076
Spring 1st	NEE13	-19.790	7.105	-2.785	0.006
Spring 1st	NEE15	-16.737	7.196	-2.326	0.021
Spring 1st	NEE50	20.148	8.265	2.438	0.015
Spring 1st	NEE81	-3.599	7.424	-0.485	0.628
Spring 1st	NEE82	37.150	7.130	5.210	0.000
Spring 1st	NEE83	-9.465	5.217	-1.814	0.070
Spring 1st	NEE84	-13.333	8.721	-1.529	0.127
Spring 1st	EDAVG	1.095	0.444	2.463	0.014
Spring 1st	FEMALE	-4.915	2.586	-1.901	0.058
Growth Rate 1-3	BASE	47.212	3.472	13.599	0.000
Growth Rate 1-3	PGMIS	-4.025	2.234	-1.802	0.072
Growth Rate 1-3	PONE13	-3.354	7.123	-0.471	0.638
Growth Rate 1-3	NIS01	5.097	2.445	2.085	0.038
Growth Rate 1-3	NIS10	-2.013	5.885	-0.342	0.733
Growth Rate 1-3	NIS7A	9.902	3.591	2.758	0.006
Growth Rate 1-3	NIS7B	2.237	3.672	0.609	0.543
Growth Rate 1-3	NIS7C	-2.858	4.565	-0.626	0.532
Growth Rate 1-3	NIS7D	-0.558	4.571	-0.122	0.903
Growth Rate 1-3	NEE02	-0.664	3.163	-0.210	0.834
Growth Rate 1-3	NEE04	13.692	4.620	2.964	0.003
Growth Rate 1-3	NEE05	4.296	3.076	1.397	0.163
Growth Rate 1-3	NEE11	-17.329	4.123	-4.203	0.000
Growth Rate 1-3	NEE13	12.311	4.601	2.676	0.008
Growth Rate 1-3	NEE15	-19.123	4.271	-4.477	0.000
Growth Rate 1-3	NEE50	-2.938	5.162	-0.569	0.570
Growth Rate 1-3	NEE81	1.100	7.310	0.150	0.880
Growth Rate 1-3	NEE82	-10.705	5.468	-1.958	0.051
Growth Rate 1-3	NEE83	15.697	4.368	3.594	0.000
Growth Rate 1-3	NEE84	10.402	5.270	1.974	0.049
Growth Rate 1-3	FEMALE	5.924	1.735	3.415	0.001

Table 177 (continued)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Curvature 1-3	BASE	-5.565	4.604	-1.209	0.227
Curvature 1-3	PGMIS	-3.807	3.078	-1.237	0.217
Curvature 1-3	PONE13	12.241	9.748	1.256	0.210
Curvature 1-3	NIS01	-3.448	3.493	-0.987	0.324
Curvature 1-3	NIS10	0.039	7.999	0.005	0.996
Curvature 1-3	NIS7A	1.238	5.019	0.247	0.805
Curvature 1-3	NIS7B	-1.892	5.209	-0.363	0.717
Curvature 1-3	NIS7C	6.719	6.622	1.015	0.311
Curvature 1-3	NIS7D	-10.834	6.352	-1.706	0.089
Curvature 1-3	NEE02	-2.486	4.450	-0.559	0.577
Curvature 1-3	NEE04	-11.505	6.748	-1.705	0.089
Curvature 1-3	NEE05	-0.852	4.462	-0.191	0.849
Curvature 1-3	NEE11	-8.923	5.759	-1.549	0.122
Curvature 1-3	NEE13	17.682	6.319	2.798	0.005
Curvature 1-3	NEE15	-5.963	5.920	-1.007	0.314
Curvature 1-3	NEE50	-0.372	7.321	-0.051	0.960
Curvature 1-3	NFE81	6.050	8.112	0.746	0.456
Curvature 1-3	NEE82	11.360	7.248	1.567	0.118
Curvature 1-3	NEE83	-7.388	5.091	-1.451	0.148
Curvature 1-3	NEE84	-1.578	8.314	-0.190	0.850
Year-to-Year Deviations	ABSENT	-0.312	0.124	-2.511	0.012

ERIC

Table 178

One-Program Schools 1-3 Analysis
School Curvature Model for Language

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	229.705	12.633	18.183	0.000
Spring 1st	PGMIS	3.875	5.052	0.767	0.444
Spring 1st	PONE13	29.529	18.252	1.618	0.107
Spring 1st	NIS01	1.659	5.945	0.279	0.780
Spring 1st	NIS10	16.969	14.521	1.169	0.243
Spring 1st	NIS7A	25.265	8.002	3.158	0.002
Spring 1st	NIS7B	-7.946	7.552	-1.052	0.293
Spring 1st	NIS7C	-13.646	11.682	-1.168	0.243
Spring 1st	NIS7D	2.714	9.595	0.283	0.777
Spring 1st	NEE02	45.492	7.513	6.055	0.000
Spring 1st	NEE04	28.193	11.894	2.370	0.018
Spring 1st	NEE05	16.323	7.678	2.126	0.034
Spring 1st	NEE11	-12.325	10.607	-1.162	0.246
Spring 1st	NEE13	-8.040	11.198	-0.718	0.473
Spring 1st	NEE15	-14.016	11.376	-1.232	0.219
Spring 1st	NEE50	10.620	13.050	0.814	0.416
Spring 1st	NEE81	-10.237	11.769	-0.870	0.385
Spring 1st	NEE82	22.060	11.281	1.955	0.051
Spring 1st	NEE83	-11.191	8.264	-1.354	0.177
Spring 1st	NEE84	-39.071	13.739	-2.844	0.005
Spring 1st	EDAVG	1.752	0.721	2.431	0.016
Spring 1st	FEMALE	3.359	4.072	0.825	0.410
Growth Rate 1-3	BASE	70.452	6.122	11.507	0.000
Growth Rate 1-3	PGMIS	4.251	3.947	1.077	0.282
Growth Rate 1-3	PONE13	-33.405	12.568	-2.658	0.008
Growth Rate 1-3	NIS01	3.254	4.335	0.751	0.453
Growth Rate 1-3	NIS10	-14.900	10.448	-1.426	0.155
Growth Rate 1-3	NIS7A	27.086	6.325	4.283	0.000
Growth Rate 1-3	NIS7B	18.521	6.483	2.857	0.005
Growth Rate 1-3	NIS7C	-10.685	8.096	-1.320	0.188
Growth Rate 1-3	NIS7D	-21.736	8.072	-2.693	0.007
Growth Rate 1-3	NEE02	-5.889	5.591	-1.053	0.293
Growth Rate 1-3	NEE04	9.852	8.204	1.201	0.231
Growth Rate 1-3	NEE05	5.730	5.449	1.052	0.294
Growth Rate 1-3	NEE11	-21.722	7.314	-2.970	0.003
Growth Rate 1-3	NEE13	-5.203	8.151	-0.638	0.524
Growth Rate 1-3	NEE15	-23.853	7.586	-3.145	0.002
Growth Rate 1-3	NEE50	12.918	9.135	1.414	0.158
Growth Rate 1-3	NEE81	1.012	12.772	0.079	0.937
Growth Rate 1-3	NEE82	8.781	9.619	0.913	0.362
Growth Rate 1-3	NEE83	14.312	7.642	1.873	0.062
Growth Rate 1-3	NEE84	19.397	9.296	2.087	0.038
Growth Rate 1-3	FEMALE	14.131	3.059	4.620	0.000

Table 178 (continued)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Curvature 1-3	BASE	-3.423	7.537	-0.454	0.650
Curvature 1-3	PGMIS	-10.328	5.043	-2.048	0.041
Curvature 1-3	PONE13	-3.853	15.943	-0.242	0.809
Curvature 1-3	NIS01	-0.669	5.689	-0.118	0.906
Curvature 1-3	NIS10	-0.729	13.025	-0.056	0.955
Curvature 1-3	NIS7A	9.491	8.274	1.147	0.252
Curvature 1-3	NIS7B	-8.380	8.739	-0.959	0.338
Curvature 1-3	NIS7C	28.547	10.747	2.656	0.008
Curvature 1-3	NIS7D	-29.516	10.577	-2.791	0.006
Curvature 1-3	NEE02	3.170	7.235	0.438	0.661
Curvature 1-3	NEE04	3.359	10.907	0.308	0.758
Curvature 1-3	NEE05	-9.903	7.230	-1.370	0.172
Curvature 1-3	NEE11	-14.446	9.363	-1.543	0.124
Curvature 1-3	NEE13	-13.747	10.267	-1.339	0.181
Curvature 1-3	NEE15	2.770	9.614	0.288	0.773
Curvature 1-3	NEE50	30.826	11.908	2.589	0.010
Curvature 1-3	NEE81	-7.519	13.596	-0.553	0.581
Curvature 1-3	NEE82	-7.887	12.075	-0.653	0.514
Curvature 1-3	NEE83	1.629	8.468	0.192	0.848
Curvature 1-3	NEE84	3.470	13.411	0.259	0.796
Year-to-Year Deviations	ABSENT	-0.087	0.203	-0.426	0.670

Table 179

One-Program Schools 1-3 Analysis:
School Curvature Model for Reading

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	226.108	10.325	21.899	0.000
Spring 1st	PGMIS	-5.960	4.109	-1.451	0.148
Spring 1st	PONE13	6.964	14.864	0.469	0.640
Spring 1st	NIS01	0.657	4.825	0.136	0.892
Spring 1st	NIS10	19.939	11.763	1.695	0.091
Spring 1st	NIS7A	28.218	6.541	4.314	0.000
Spring 1st	NIS7B	0.637	6.179	0.103	0.918
Spring 1st	NIS7C	-27.985	9.512	-2.942	0.003
Spring 1st	NIS7D	-1.101	7.839	-0.140	0.888
Spring 1st	NEE02	33.005	6.098	5.412	0.000
Spring 1st	NEE04	5.765	9.637	0.598	0.550
Spring 1st	NEE05	10.147	6.232	1.628	0.104
Spring 1st	NEE11	-10.421	8.597	-1.212	0.226
Spring 1st	NEE13	-0.096	9.062	-0.011	0.992
Spring 1st	NEE15	-7.133	9.225	-0.773	0.440
Spring 1st	NEE50	-2.424	10.607	-0.229	0.819
Spring 1st	NEE81	0.416	9.607	0.043	0.966
Spring 1st	NEE82	13.274	9.203	1.442	0.150
Spring 1st	NEE83	-7.393	6.736	-1.098	0.273
Spring 1st	NEE84	-21.881	11.205	-1.953	0.052
Spring 1st	EDAVG	1.485	0.590	2.516	0.012
Spring 1st	FEMALE	3.983	3.307	1.204	0.229
Growth Rate 1-3	BASE	59.864	5.315	11.263	0.000
Growth Rate 1-3	PGMIS	2.994	3.424	0.874	0.382
Growth Rate 1-3	PONE13	-13.986	10.907	-1.282	0.201
Growth Rate 1-3	NIS01	1.753	3.752	0.467	0.641
Growth Rate 1-3	NIS10	-6.422	9.031	-0.711	0.477
Growth Rate 1-3	NIS7A	16.271	5.501	2.958	0.003
Growth Rate 1-3	NIS7B	0.223	5.646	0.040	0.968
Growth Rate 1-3	NIS7C	-2.692	6.999	-0.385	0.701
Growth Rate 1-3	NIS7D	-20.165	7.018	-2.873	0.004
Growth Rate 1-3	NEE02	-4.705	4.845	-0.971	0.332
Growth Rate 1-3	NEE04	13.575	7.087	1.916	0.056
Growth Rate 1-3	NEE05	1.993	4.715	0.423	0.673
Growth Rate 1-3	NEE11	-15.018	6.325	-2.374	0.018
Growth Rate 1-3	NEE13	4.257	7.046	0.604	0.546
Growth Rate 1-3	NEE15	-19.864	6.555	-3.030	0.003
Growth Rate 1-3	NEE50	8.406	7.910	1.063	0.289
Growth Rate 1-3	NEE81	-3.207	11.144	-0.288	0.774
Growth Rate 1-3	NEE82	10.211	8.375	1.219	0.224
Growth Rate 1-3	NEE83	11.000	6.661	1.651	0.099
Growth Rate 1-3	NEE84	5.275	8.060	0.654	0.513
Growth Rate 1-3	FEMALE	4.149	2.655	1.562	0.119

Table 179 (continued)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Curvature 1-3	BASE	-12.101	6.846	-1.767	0.078
Curvature 1-3	PGMIS	-3.525	4.582	-0.769	0.442
Curvature 1-3	PONE13	9.811	14.498	0.677	0.499
Curvature 1-3	NIS01	4.096	5.193	0.789	0.431
Curvature 1-3	NIS10	-13.671	11.886	-1.150	0.251
Curvature 1-3	NIS7A	9.267	7.509	1.234	0.218
Curvature 1-3	NIS7B	-15.098	7.869	-1.919	0.056
Curvature 1-3	NIS7C	24.118	9.820	2.456	0.014
Curvature 1-3	NIS7D	-3.378	9.546	-0.354	0.724
Curvature 1-3	NEE02	3.722	6.605	0.564	0.573
Curvature 1-3	NEE04	10.884	9.999	1.088	0.277
Curvature 1-3	NEE05	3.720	6.615	0.562	0.574
Curvature 1-3	NEE11	-1.120	8.547	-0.131	0.896
Curvature 1-3	NEE13	-15.340	9.379	-1.636	0.103
Curvature 1-3	NEE15	-12.177	8.780	-1.387	0.166
Curvature 1-3	NEE50	4.326	10.868	0.398	0.691
Curvature 1-3	NEE81	-26.056	12.060	-2.161	0.031
Curvature 1-3	NEE82	6.669	10.776	0.619	0.536
Curvature 1-3	NEE83	-0.274	7.551	-0.036	0.971
Curvature 1-3	NEE84	16.358	12.305	1.329	0.185
Year-to-Year Deviations	ABSENT	-0.097	0.174	-0.559	0.576

For all three subtests, the overall school-to-school variability in initial status and growth rate is highly significant (p less than .0005). In contrast, the school-to-school variability in curvature is never significant at the .01 level. For the mathematics and language subtests, the school-to-school variability in curvature just achieves statistical significance at the .05 level (p = .040 and p = .048, respectively). For the reading subtest, the school-to-school variability is not significant (p = .095).

The inclusion of school as a predictor of curvature complicates the interpretation of the results (see Table 179). Most of the school-to-school differences in curvature were not significant at the .05 level. Of the school-to-school differences that may be significant (i.e., NIS7C and NEE81), the apparent school differences in curvature may be a result of the combination of district H schools into groups based on preliminary analyses. For all of these reasons, the two-level model without school curvature was chosen as the basic model.

<u>How does modeling school curvature affect the results for the three subtests?</u>

Allowing school curvature in the model changes the results for the mathematics subtest very little. A comparison of Table 174 with Table 177 reveals that most parameters barely change. There are still no statistically significant program differences in initial status, growth rate, or curvature after adjusting for propensity. However, the program difference in growth rate is increased slightly, from -3.188 to -4.025, and approaches statistical significance (p = .072). The estimated program differences at each grade remain small. At spring of first grade, the estimated program difference is 2.318, at spring of second grade it is 2.100, and at spring of third grade it is -5.732. Although the mathematics growth curves are not significantly different between the two programs, it is intriguing that the pattern of scores is familiar from the language and reading analyses (both for two-program schools and for one-program schools): the immersion strategy students have a temporary advantage in test scores relative to early-exit students followed by flattening growth curves.

For the language subtest, the most important change resulting from including school as a predictor of curvature in the model for the language subtest is that the curvature difference between the programs is reduced. Instead of a difference of -14.069 (p = .001), the difference is reduced to -10.328 (p = .041). However, the pattern of program differences is still the same: at first grade, the difference is only 0.432, at second grade the difference is 15.011, and at third grade the difference has dropped to 8.934. Also, the propensity score (PONE13) as a predictor of curvature is considerably weakened, indicating that the propensity score was accounting for some of the school-to-school differences in curvature. For the language subtest, the basic findings are unaffected by permitting schools to have different curvature: the immersion strategy students have higher scores than early-exit students at second grade but lower growth from second to third grade.

For the reading subtest, the program effects are almost unchanged by allowing curvature to vary by school. The initial status, growth rate, and curvature are not significantly different for the two programs. The estimated program difference at first grade changes from -7.269 to -7.135, the estimated second grade difference changes from 0.100 to -0.616, and the estimated third grade difference changes from -0.767 to -1.147. The effect of the propensity score as a predictor of curvature is lessened by the inclusion of school-level curvature for the reading subtest, just as it was for the language subtest, again indicating that some of the school-to-school differences in curvature are accommodated by the propensity score.

<u>Summary</u>

Allowing curvature to vary by school changes the results of the basic 1-3 analyses very little. The conclusions from the basic two-level models are not altered in the more complicated models that include school differences in curvature.

Including the Pretest as a Predictor in the 1-3 Analyses

What was the purpose of the 1-3 pretest analyses?

The basic 1-3 analyses of one-program schools does not include the fall kindergarten test scores as a predictor. Just as for the two-program schools, many students did not have pretest scores available. In order to estimate program differences from first through third grade using as many students as possible, pretest scores were not used in the basic 1-3 analyses. The main purpose of the 1-3 pretest analyses was to evaluate the sensitivity of estimated program differences to the inclusion of the pretest as a predictor of propensity or as a predictor of the parameters of the individual growth curves.

How was the effect of the pretest evaluated?

The effect of pretest on the 1-3 analyses was evaluated in a series of four analyses for each subtest. First, the basic model was estimated again, but using only those students with pretest scores available. This permits an assessment of how the results change with the reduced sample. Second, the pretest was included as a predictor of the initial status but with the original propensity score. Third, the basic model was estimated using students with pretest scores but with the propensity score developed for those students. The pretest was not included as a predictor of propensity nor as a predictor of initial status in this analysis. Fourth, pretest was included as a predictor of initial status and as a predictor of propensity.

Several of the one-program schools and school groups had no students with pretest and so had to be omitted from the pretest analyses. Among the immersion strategy schools, six of the twelve district P schools had no students with pretest scores but only the IS7C school group had to be dropped. Among the early-exit schools, two had to be dropped: EE50, the only school in district F, and EE84, one of four schools in district I. Although early-exit school 4 (EE04) in district A had only two students with pretest scores available, it was included in the pretest analyses. The other schools and school groups included in the pretest analyses had 8 to 34 students.

What were the results of the 1-3 pretest analyses for the three subtests?

Table 180 shows the basic model estimated for those students with pretest scores available and should be compared with Table 174. Although the propensity-adjusted program effects are still nonsignificant, the estimated program difference for the growth rate changes from -3.188 to 8.264 and approaches statistical significance ($p = .061$). This change indicates that among the students (and schools) with pretest scores available, the students in IS schools may have higher mathematics growth rates than the students in EE schools. The estimated program differences at first, second, and third grade indicate this change from the basic model estimated using all students; the differences are estimated to be -3.274, 9.093, 13.254, respectively.

ERIC

Table 180

One-Program Schools 1-3 Analysis:
Two-Level Model for Mathematics (Students with Pretest)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	254.685	9.253	27.524	0.000
Spring 1st	PGMIS	-1.906	4.575	-0.417	0.677
Spring 1st	PONE13	-8.254	14.299	-0.577	0.564
Spring 1st	NIS01	-9.194	4.621	-1.990	0.048
Spring 1st	NIS10	1.213	9.453	0.128	0.898
Spring 1st	NIS7A	10.465	8.146	1.285	0.200
Spring 1st	NIS7B	4.031	6.142	0.656	0.512
Spring 1st	NIS7D	-18.423	8.103	-2.274	0.024
Spring 1st	NEE02	25.455	5.664	4.494	0.000
Spring 1st	NEE04	18.453	16.489	1.119	0.264
Spring 1st	NEE05	-2.553	5.932	-0.430	0.667
Spring 1st	NEE11	-11.007	6.690	-1.645	0.101
Spring 1st	NEE13	-19.187	7.065	-2.716	0.007
Spring 1st	NEE15	-16.177	7.184	-2.252	0.025
Spring 1st	NEE81	-0.140	9.006	-0.016	0.988
Spring 1st	NEE82	24.639	9.140	2.696	0.008
Spring 1st	NEE83	-14.123	6.438	-2.194	0.029
Spring 1st	EDAVG	1.134	0.580	1.955	0.052
Spring 1st	FEMALE	-6.611	3.191	-2.072	0.039
Growth Rate 1-3	BASE	45.558	4.206	10.831	0.000
Growth Rate 1-3	PGMIS	8.264	4.390	1.882	0.061
Growth Rate 1-3	PONE13	-10.617	9.945	-1.068	0.287
Growth Rate 1-3	NIS01	-1.807	3.980	-0.454	0.650
Growth Rate 1-3	NIS10	-13.042	6.748	-1.933	0.054
Growth Rate 1-3	NIS7A	20.044	8.430	2.378	0.018
Growth Rate 1-3	NIS7B	1.483	5.988	0.248	0.805
Growth Rate 1-3	NIS7D	15.168	8.148	1.862	0.064
Growth Rate 1-3	NEE02	1.006	4.208	0.239	0.811
Growth Rate 1-3	NEE04	-6.541	12.589	-0.520	0.604
Growth Rate 1-3	NEE05	11.715	4.337	2.701	0.007
Growth Rate 1-3	NEE11	-16.061	4.491	-3.576	0.000
Growth Rate 1-3	NEE13	10.904	5.026	2.170	0.031
Growth Rate 1-3	NEE15	-17.369	4.617	-3.762	0.000
Growth Rate 1-3	NEE81	0.677	9.096	0.074	0.941
Growth Rate 1-3	NEE82	-5.613	8.888	-0.631	0.528
Growth Rate 1-3	NEE83	26.765	6.159	4.346	0.000
Growth Rate 1-3	FEMALE	9.042	2.435	3.713	0.000
Curvature 1-3	BASE	-5.866	3.473	-1.689	0.093
Curvature 1-3	PGMIS	-4.103	4.319	-0.950	0.343
Curvature 1-3	PONE13	12.754	9.758	1.307	0.192
Year-to-Year Deviations	ABSENT	-0.294	0.156	-1.886	0.061

Table 181 shows the model after adding the mathematics pretest score (ANALPREM) as a predictor of initial status. The resulting model is very similar to the one shown in Table 180. Both of these models use the propensity score developed for the 1-3 analyses that use all students. Two additional propensity scores were calculated to reflect the reduced sample of students who had pretest scores available. The first of these, PONE13PN (one-program school, grades 1-3, with pretest, and pretest not used calculating propensity score), used the same variables as the original propensity score (PONE13) but the estimated coefficients were allowed to change. The second, PONE13PP (one-program school, grades 1-3, with pretest, and pretest used to calculate propensity score), included the pretest as a predictor of propensity as well as the variables used in the other two propensity scores. Table 182 shows the model for mathematics using the PONE13PN propensity score; in this model, the pretest is not used as a predictor of either the propensity score or the initial status. In the model given in Table 183, both the PONE13PP propensity score and the mathematics pretest are included; the pretest is used both as a predictor of the propensity score and as a predictor of initial status.

Table 181

One-Program Schools 1-3 Analysis:
Two-Level Model for Mathematics (Pretest as Predictor)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	152.976	17.114	8.938	0.000
Spring 1st	PGMIS	-0.306	4.281	-0.071	0.943
Spring 1st	PONE13	3.155	13.379	0.236	0.814
Spring 1st	NIS01	-14.112	4.365	-3.233	0.001
Spring 1st	NIS10	9.587	8.860	1.082	0.280
Spring 1st	NIS7A	8.743	7.671	1.140	0.256
Spring 1st	NIS7B	3.503	5.760	0.608	0.544
Spring 1st	NIS7D	-21.777	7.612	-2.861	0.005
Spring 1st	NEE02	15.654	5.452	2.871	0.004
Spring 1st	NEE04	15.301	15.353	0.997	0.320
Spring 1st	NEE05	-4.458	5.523	-0.807	0.420
Spring 1st	NEE11	-5.828	6.244	-0.933	0.352
Spring 1st	NEE13	-11.168	6.652	-1.679	0.095
Spring 1st	NEE15	-8.475	6.745	-1.256	0.210
Spring 1st	NEE81	6.440	8.425	0.764	0.445
Spring 1st	NEE82	16.402	8.578	1.912	0.057
Spring 1st	NEE83	-19.203	6.034	-3.182	0.002
Spring 1st	EDAVG	1.457	0.534	2.729	0.007
Spring 1st	FEMALE	-7.034	2.967	-2.370	0.019
Spring 1st	ANALPREM	0.553	0.080	6.877	0.000
Growth Rate 1-3	BASE	45.559	4.205	10.834	0.000
Growth Rate 1-3	PGMIS	8.298	4.393	1.889	0.060
Growth Rate 1-3	PONE13	-11.034	9.937	-1.110	0.268
Growth Rate 1-3	NIS01	-1.667	3.984	-0.419	0.676
Growth Rate 1-3	NIS10	-13.089	6.744	-1.941	0.053
Growth Rate 1-3	NIS7A	19.892	8.449	2.355	0.019
Growth Rate 1-3	NIS7B	1.490	6.000	0.248	0.804
Growth Rate 1-3	NIS7D	15.450	8.162	1.893	0.060
Growth Rate 1-3	NEE02	1.295	4.207	0.308	0.759
Growth Rate 1-3	NEE04	-6.808	12.583	-0.541	0.589
Growth Rate 1-3	NEE05	11.814	4.336	2.725	0.007
Growth Rate 1-3	NEE11	-16.111	4.490	-3.589	0.000
Growth Rate 1-3	NEE13	10.785	5.018	2.149	0.033
Growth Rate 1-3	NEE15	-17.300	4.616	-3.748	0.000
Growth Rate 1-3	NEE81	0.687	9.116	0.075	0.940
Growth Rate 1-3	NEE82	-5.683	8.907	-0.638	0.524
Growth Rate 1-3	NEE83	26.884	6.171	4.357	0.000
Growth Rate 1-3	FEMALE	9.143	2.434	3.757	0.000
Curvature 1-3	BASE	-5.905	3.481	-1.696	0.091
Curvature 1-3	PGMIS	-4.098	4.329	-0.946	0.345
Curvature 1-3	PONE13	12.580	9.781	1.286	0.200
Year-to-Year Deviations	ABSENT	-0.250	0.150	-1.662	0.098

Table 182

One-Program Schools 1-3 Analysis:
Two-Level Model for Mathematics (With Propensity for Pretest Students)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	249.852	8.160	30.621	0.000
Spring 1st	PGMIS	-3.238	4.499	-0.720	0.472
Spring 1st	PONE13PN	0.583	12.579	0.046	0.963
Spring 1st	NIS01	-9.237	4.665	-1.980	0.049
Spring 1st	NIS10	4.735	8.844	0.535	0.593
Spring 1st	NIS7A	9.203	8.106	1.135	0.257
Spring 1st	NIS7B	2.793	6.026	0.464	0.643
Spring 1st	NIS7D	-20.340	7.911	-2.571	0.011
Spring 1st	NEE02	24.608	5.736	4.290	0.000
Spring 1st	NEE04	18.217	16.632	1.095	0.275
Spring 1st	NEE05	-3.103	5.984	-0.519	0.605
Spring 1st	NEE11	-8.756	6.316	-1.386	0.167
Spring 1st	NEE13	-17.406	6.801	-2.559	0.011
Spring 1st	NEE15	-13.928	6.847	-2.034	0.043
Spring 1st	NEE81	-1.721	8.840	-0.195	0.846
Spring 1st	NEE82	22.896	8.928	2.565	0.011
Spring 1st	NEE83	-15.701	6.178	-2.542	0.012
Spring 1st	EDAVG	1.355	0.554	2.446	0.015
Spring 1st	FEMALE	-6.758	3.216	-2.101	0.037
Growth Rate 1-3	BASE	44.620	3.787	11.783	0.000
Growth Rate 1-3	PGMIS	8.099	4.327	1.872	0.062
Growth Rate 1-3	PONE13PN	-9.524	9.004	-1.058	0.291
Growth Rate 1-3	NIS01	-1.286	3.934	-0.327	0.744
Growth Rate 1-3	NIS10	-11.968	6.382	-1.875	0.062
Growth Rate 1-3	NIS7A	19.564	8.385	2.333	0.020
Growth Rate 1-3	NIS7B	0.936	5.936	0.158	0.875
Growth Rate 1-3	NIS7D	14.266	7.993	1.785	0.076
Growth Rate 1-3	NEE02	1.519	4.178	0.363	0.717
Growth Rate 1-3	NEE04	-6.253	12.484	-0.501	0.617
Growth Rate 1-3	NEE05	12.055	4.273	2.821	0.005
Growth Rate 1-3	NEE11	-15.230	4.262	-3.573	0.000
Growth Rate 1-3	NEE13	11.785	4.807	2.452	0.015
Growth Rate 1-3	NEE15	-16.594	4.427	-3.748	0.000
Growth Rate 1-3	NEE81	-0.553	8.948	-0.062	0.951
Growth Rate 1-3	NEE82	-6.899	8.739	-0.789	0.431
Growth Rate 1-3	NEE83	25.713	5.998	4.287	0.000
Growth Rate 1-3	FEMALE	8.996	2.414	3.727	0.000
Curvature 1-3	BASE	-5.403	3.341	-1.617	0.107
Curvature 1-3	PGMIS	-3.920	4.323	-0.907	0.365
Curvature 1-3	PONE13PN	11.437	9.514	1.202	0.231
Year-to-Year Deviations	ABSENT	-0.288	0.156	-1.847	0.066

Table 183

One-Program Schools 1-3 Analysis:
Two-Level Model for Mathematics
(With Pretest as Predictor of Propensity and Initial Status)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	156.540	19.278	8.120	0.000
Spring 1st	PGMIS	-0.645	4.245	-0.152	0.879
Spring 1st	PONE13PP	3.408	11.139	0.306	0.760
Spring 1st	NIS01	-14.242	4.436	-3.211	0.002
Spring 1st	NIS10	9.116	8.215	1.110	0.268
Spring 1st	NIS7A	8.910	7.593	1.174	0.242
Spring 1st	NIS7B	3.683	5.579	0.660	0.510
Spring 1st	NIS7D	-21.505	7.402	-2.905	0.004
Spring 1st	NEE02	15.888	5.486	2.896	0.004
Spring 1st	NEE04	15.601	15.434	1.011	0.313
Spring 1st	NEE05	-4.742	5.611	-0.845	0.399
Spring 1st	NEE11	-6.216	5.841	-1.064	0.288
Spring 1st	NEE13	-11.698	6.416	-1.823	0.070
Spring 1st	NEE15	-9.042	6.374	-1.419	0.157
Spring 1st	NEE81	6.089	8.333	0.731	0.466
Spring 1st	NEE82	17.392	8.291	2.098	0.037
Spring 1st	NEE83	-18.420	5.653	-3.258	0.001
Spring 1st	EDAVG	1.430	0.498	2.874	0.004
Spring 1st	FEMALE	-7.050	2.977	-2.368	0.019
Spring 1st	ANALPREM	0.535	0.091	5.896	0.000
Growth Rate 1-3	BASE	46.037	3.494	13.176	0.000
Growth Rate 1-3	PGMIS	9.718	4.358	2.230	0.027
Growth Rate 1-3	PONE13PP	-16.086	8.026	-2.004	0.046
Growth Rate 1-3	NIS01	-0.546	3.928	-0.139	0.890
Growth Rate 1-3	NIS10	-13.455	6.175	-2.179	0.030
Growth Rate 1-3	NIS7A	19.852	8.380	2.369	0.019
Growth Rate 1-3	NIS7B	1.048	5.927	0.177	0.860
Growth Rate 1-3	NIS7D	15.676	7.938	1.975	0.049
Growth Rate 1-3	NEE02	1.596	4.110	0.388	0.698
Growth Rate 1-3	NEE04	-7.594	12.476	-0.609	0.543
Growth Pate 1-3	NEE05	13.193	4.282	3.081	0.002
Growth Rate 1-3	NEE11	-16.211	4.174	-3.884	0.000
Growth Rate 1-3	NEE13	10.990	4.703	2.337	0.020
Growth Rate 1-3	NEE15	-16.994	4.333	-3.922	0.000
Growth Rate 1-3	NEE81	2.075	9.104	0.228	0.820
Growth Rate 1-3	NEE82	-7.290	8.670	-0.841	0.401
Growth Rate 1-3	NEE83	25.496	5.888	4.330	0.000
Growth Rate 1-3	FEMALE	9.302	2.410	3.860	0.000
Curvature 1-3	BASE	-4.791	3.146	-1.523	0.129
Curvature 1-3	PGMIS	-3.826	4.405	-0.869	0.386
Curvature 1-3	PONE13PP	8.988	8.949	1.004	0.316
Year-to-Year Deviations	ABSENT	-0.249	0.150	-1.660	0.098

The estimated program effects differ little across these four models for the mathematics subtest. The estimated program difference in growth rate is near statistical significance for the first three models and is significant in the fourth (difference = 9.712, p = .027). The magnitude of the program difference in growth rate is fairly insensitive to the form of the model, ranging only from about 8.1 to about 9.7 ESS points per year.

The 1-3 analyses for the language subtest on the students who have pretest scores available parallel those for the mathematics subtest. Table 184 shows the basic model estimated for those students with pretest scores; it should be compared with Table 175. The estimated program difference for the growth rate is dramatically increased, from 4.736 (p = .223) to 20.222 (p = .008). The program effect on curvature is somewhat reduced, from -14.069 (p = .001) to -8.922 (p = .213).

Just as for the mathematics subtest, the estimated program differences in growth curves for the language subtest are not very sensitive to the inclusion of the pretest score itself or to changing the propensity score. Table 185, Table 186, and Table 187 all show program effects that are similar to those in the basic model for students with pretest scores available (see Table 184).

Table 184

One-Program Schools 1-3 Analysis:
Two-Level Model for Language (Students with Pretest)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	218.836	14.328	15.273	0.000
Spring 1st	PGMIS	0.020	6.928	0.003	0.998
Spring 1st	PONE13	47.237	21.760	2.171	0.031
Spring 1st	NIS01	8.467	6.952	1.218	0.225
Spring 1st	NIS10	26.248	14.278	1.838	0.067
Spring 1st	NIS7A	21.096	12.418	1.699	0.091
Spring 1st	NIS7B	-11.690	9.367	-1.248	0.213
Spring 1st	NIS7D	-22.787	12.283	-1.855	0.065
Spring 1st	NEE02	57.983	8.515	6.809	0.000
Spring 1st	NEE04	13.818	24.803	0.557	0.578
Spring 1st	NEE05	-5.029	8.914	-0.564	0.573
Spring 1st	NEE11	-7.017	10.102	-0.695	0.488
Spring 1st	NEE13	-5.153	10.612	-0.486	0.628
Spring 1st	NEE15	-8.430	10.867	-0.776	0.439
Spring 1st	NEE81	-17.173	13.547	-1.268	0.206
Spring 1st	NEE82	9.269	13.768	0.673	0.501
Spring 1st	NEE83	-13.157	9.728	-1.352	0.178
Spring 1st	EDAVG	2.151	0.912	2.358	0.019
Spring 1st	FEMALE	3.813	4.794	0.795	0.427
Growth Rate 1-3	BASE	71.063	7.480	9.501	0.000
Growth Rate 1-3	PGMIS	20.222	7.610	2.657	0.008
Growth Rate 1-3	PONE13	-49.725	17.814	-2.791	0.006
Growth Rate 1-3	NIS01	-12.009	6.930	-1.733	0.084
Growth Rate 1-3	NIS10	-30.545	12.026	-2.540	0.012
Growth Rate 1-3	NIS7A	9.442	14.226	0.664	0.508
Growth Rate 1-3	NIS7B	30.205	10.129	2.982	0.003
Growth Rate 1-3	NIS7D	13.173	13.825	0.953	0.342
Growth Rate 1-3	NEE02	-4.706	7.504	-0.627	0.531
Growth Rate 1-3	NEE04	-16.010	22.542	-0.710	0.478
Growth Rate 1-3	NEE05	24.030	7.751	3.100	0.002
Growth Rate 1-3	NEE11	-21.686	8.051	-2.694	0.008
Growth Rate 1-3	NEE13	-4.743	9.001	-0.527	0.599
Growth Rate 1-3	NEE15	-23.010	8.296	-2.774	0.006
Growth Rate 1-3	NEE81	32.545	15.479	2.103	0.037
Growth Rate 1-3	NEE82	18.453	15.151	1.218	0.224
Growth Rate 1-3	NEE83	11.075	10.492	1.055	0.292
Growth Rate 1-3	FEMALE	17.125	4.364	3.924	0.000
Curvature 1-3	BASE	-8.964	5.725	-1.566	0.119
Curvature 1-3	PGMIS	-8.922	7.142	-1.249	0.213
Curvature 1-3	PONE13	0.415	16.153	0.026	0.980
Year-to-Year Deviations	ABSENT	-0.049	0.248	-0.198	0.843

ERIC

Table 185

One-Program Schools 1-3 Analysis:
Two-Level Model for Language (Pretest as Predictor)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	105.758	24.060	4.396	0.000
Spring 1st	PGMIS	4.285	6.638	0.646	0.519
Spring 1st	PONE13	51.350	20.621	2.490	0.013
Spring 1st	NIS01	4.423	6.639	0.666	0.506
Spring 1st	NIS10	33.160	13.568	2.444	0.015
Spring 1st	NIS7A	17.553	11.900	1.475	0.142
Spring 1st	NIS7B	-14.136	8.947	-1.580	0.115
Spring 1st	NIS7D	-21.416	11.718	-1.828	0.069
Spring 1st	NEE02	45.776	8.350	5.482	0.000
Spring 1st	NEE04	6.099	23.571	0.259	0.796
Spring 1st	NEE05	-6.061	8.452	-0.717	0.474
Spring 1st	NEE11	-1.374	9.599	-0.143	0.886
Spring 1st	NEE13	2.191	10.114	0.217	0.829
Spring 1st	NEE15	-1.415	10.341	-0.137	0.891
Spring 1st	NEE81	-9.390	12.907	-0.728	0.468
Spring 1st	NEE82	3.608	13.078	0.276	0.783
Spring 1st	NEE83	-18.722	9.272	-2.019	0.045
Spring 1st	EDAVG	2.270	0.857	2.649	0.009
Spring 1st	FEMALE	2.718	4.548	0.598	0.551
Spring 1st	ANALTOBE	0.309	0.055	5.672	0.000
Growth Rate 1-3	BASE	71.208	7.532	9.454	0.000
Growth Rate 1-3	PGMIS	20.287	7.649	2.652	0.009
Growth Rate 1-3	PONE13	-50.063	17.948	-2.789	0.006
Growth Rate 1-3	NIS01	-12.026	6.969	-1.726	0.086
Growth Rate 1-3	NIS10	-30.919	12.110	-2.553	0.011
Growth Rate 1-3	NIS7A	9.421	14.274	0.660	0.510
Growth Rate 1-3	NIS7B	30.103	10.164	2.962	0.003
Growth Rate 1-3	NIS7D	13.425	13.874	0.968	0.334
Growth Rate 1-3	NEE02	-4.626	7.558	-0.612	0.541
Growth Rate 1-3	NEE04	-15.643	22.710	-0.689	0.492
Growth Rate 1-3	NEE05	24.074	7.808	3.083	0.002
Growth Rate 1-3	NEE11	-21.726	8.112	-2.678	0.008
Growth Rate 1-3	NEE13	-5.095	9.064	-0.562	0.575
Growth Rate 1-3	NEE15	-23.103	8.361	-2.763	0.006
Growth Rate 1-3	NEE81	32.564	15.542	2.095	0.037
Growth Rate 1-3	NEE82	18.430	15.215	1.211	0.227
Growth Rate 1-3	NEE83	11.157	10.535	1.059	0.291
Growth Rate 1-3	FEMALE	17.196	4.397	3.911	0.000
Curvature 1-3	BASE	-8.991	5.733	-1.568	0.118
Curvature 1-3	PGMIS	-8.841	7.154	-1.236	0.218
Curvature 1-3	PONE13	0.520	16.181	0.032	0.974
Year-to-Year Deviations	ABSENT	-0.022	0.241	-0.089	0.929

ERIC

Table 186

One-Program Schools 1-3 Analysis:
Two-Level Model for Language (With Propensity for Pretest Students)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	219.550	12.458	17.623	0.000
Spring 1st	PGMIS	-0.392	6.721	-0.058	0.954
Spring 1st	PONE13PN	49.293	18.827	2.618	0.009
Spring 1st	NIS01	7.220	6.921	1.043	0.298
Spring 1st	NIS10	25.187	13.135	1.918	0.056
Spring 1st	NIS7A	21.550	12.206	1.765	0.079
Spring 1st	NIS7B	-10.805	9.066	-1.192	0.235
Spring 1st	NIS7D	-22.175	11.834	-1.874	0.062
Spring 1st	NEE02	56.675	8.499	6.668	0.000
Spring 1st	NEE04	13.503	24.671	0.547	0.585
Spring 1st	NEE05	-5.672	8.864	-0.640	0.523
Spring 1st	NEE11	-7.928	9.375	-0.846	0.399
Spring 1st	NEE13	-6.495	10.046	-0.647	0.519
Spring 1st	NEE15	-9.425	10.186	-0.925	0.356
Spring 1st	NEE81	-15.555	13.090	-1.188	0.236
Spring 1st	NEE82	11.068	13.235	0.836	0.404
Spring 1st	NEE83	-11.783	9.184	-1.283	0.201
Spring 1st	EDAVG	2.148	0.860	2.497	0.013
Spring 1st	FEMALE	3.800	4.764	0.798	0.426
Growth Rate 1-3	BASE	67.211	6.818	9.858	0.000
Growth Rate 1-3	PGMIS	19.135	7.562	2.530	0.012
Growth Rate 1-3	PONE13PN	-42.615	16.381	-2.601	0.010
Growth Rate 1-3	NIS01	-10.750	6.925	-1.552	0.122
Growth Rate 1-3	NIS10	-25.988	11.528	-2.254	0.025
Growth Rate 1-3	NIS7A	7.855	14.211	0.553	0.581
Growth Rate 1-3	NIS7B	28.595	10.088	2.835	0.005
Growth Rate 1-3	NIS7D	9.900	13.623	0.727	0.468
Growth Rate 1-3	NEE02	-3.959	7.559	-0.524	0.601
Growth Rate 1-3	NEE04	-15.365	22.672	-0.678	0.499
Growth Rate 1-3	NEE05	24.286	7.753	3.133	0.002
Growth Rate 1-3	NEE11	-18.662	7.758	-2.405	0.017
Growth Rate 1-3	NEE13	-1.697	8.743	-0.194	0.846
Growth Rate 1-3	NEE15	-20.166	8.080	-2.496	0.013
Growth Rate 1-3	NEE81	29.177	15.312	1.906	0.058
Growth Rate 1-3	NEE82	15.027	14.985	1.003	0.317
Growth Rate 1-3	NEE83	8.120	10.276	0.790	0.430
Growth Rate 1-3	FEMALE	16.986	4.388	3.871	0.000
Curvature 1-3	BASE	-8.129	5.517	-1.473	0.142
Curvature 1-3	PGMIS	-8.200	7.161	-1.145	0.253
Curvature 1-3	PONE13PN	-3.137	15.787	-0.199	0.843
Year-to-Year Deviations	ABSENT	-0.047	0.247	-0.191	0.849

ERIC

Table 187

One-Program Schools 1-3 Analysis:
Two-Level Model for Language
(With Pretest as Predictor of Propensity and Initial Status)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	89.448	29.298	3.053	0.003
Spring 1st	PGMIS	4.551	6.566	0.693	0.489
Spring 1st	PONE13PP	45.298	17.873	2.534	0.012
Spring 1st	NIS01	2.131	6.748	0.316	0.752
Spring 1st	NIS10	29.689	12.832	2.314	0.022
Spring 1st	NIS7A	18.574	11.870	1.565	0.119
Spring 1st	NIS7B	-12.139	8.745	-1.388	0.166
Spring 1st	NIS7D	-19.048	11.378	-1.674	0.095
Spring 1st	NEE02	45.022	8.523	5.283	0.000
Spring 1st	NEE04	5.896	23.764	0.248	0.804
Spring 1st	NEE05	-7.696	8.613	-0.894	0.372
Spring 1st	NEE11	-3.480	9.197	-0.378	0.705
Spring 1st	NEE13	0.173	9.884	0.018	0.986
Spring 1st	NEE15	-3.992	9.923	-0.402	0.688
Spring 1st	NEE81	-8.335	12.797	-0.651	0.515
Spring 1st	NEE82	7.487	12.684	0.590	0.556
Spring 1st	NEE83	-15.460	8.796	-1.758	0.080
Spring 1st	EDAVG	2.075	0.808	2.569	0.011
Spring 1st	FEMALE	2.857	4.582	0.624	0.534
Spring 1st	ANALTOBE	0.370	0.064	5.761	0.000
Growth Rate 1-3	BASE	66.308	6.254	10.602	0.000
Growth Rate 1-3	PGMIS	20.224	7.594	2.663	0.008
Growth Rate 1-3	PONE13PP	-40.941	14.547	-2.814	0.005
Growth Rate 1-3	NIS01	-10.704	6.885	-1.555	0.121
Growth Rate 1-3	NIS10	-25.289	11.106	-2.277	0.024
Growth Rate 1-3	NIS7A	7.765	14.136	0.549	0.583
Growth Rate 1-3	NIS7B	28.085	10.025	2.802	0.006
Growth Rate 1-3	NIS7D	10.603	13.462	0.788	0.432
Growth Rate 1-3	NEE02	-6.084	7.400	-0.822	0.412
Growth Rate 1-3	NEE04	-16.734	22.555	-0.742	0.459
Growth Rate 1-3	NEE05	24.963	7.731	3.229	0.001
Growth Rate 1-3	NEE11	-18.386	7.560	-2.432	0.016
Growth Rate 1-3	NEE13	-1.272	8.522	-0.149	0.881
Growth Rate 1-3	NEE15	-19.266	7.868	-2.449	0.015
Growth Rate 1-3	NEE81	32.767	15.503	2.114	0.036
Growth Rate 1-3	NEE82	13.875	14.795	0.938	0.349
Growth Rate 1-3	NEE83	7.100	10.042	0.707	0.480
Growth Rate 1-3	FEMALE	16.979	4.361	3.894	0.000
Curvature 1-3	BASE	-9.266	5.181	-1.788	0.075
Curvature 1-3	PGMIS	-8.872	7.278	-1.219	0.224
Curvature 1-3	PONE13PP	0.621	14.840	0.042	0.967
Year-to-Year Deviations	ABSENT	-0.025	0.242	-0.105	0.917

For the reading subtest, unlike the mathematics and language subtests, restricting the analysis to those students who have pretest scores available has relatively little effect. Table 188 shows only minor changes in the estimated program effects compared with Table 176. The programs exhibit no statistically significant differences in initial status, growth rate, or curvature in either analysis. Like the mathematics and language subtests, the reading subtest is insensitive to the form of the analysis on the students with pretest scores available (see Table 189, Table 190, and Table 191). The estimated program effects show almost no deviations across the alternative analyses.

ERIC

Table 188

One-Program Schools 1-3 Analysis:
Two-Level Model for Reading (Students with Pretest)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	207.785	13.422	15.481	0.000
Spring 1st	PGMIS	-1.733	6.593	-0.263	0.793
Spring 1st	PONE13	22.201	20.629	1.076	0.283
Spring 1st	NIS01	3.369	6.654	0.506	0.613
Spring 1st	NIS10	23.562	13.618	1.730	0.085
Spring 1st	NIS7A	25.928	11.782	2.201	0.029
Spring 1st	NIS7B	-9.150	8.882	-1.030	0.304
Spring 1st	NIS7D	-22.313	11.698	-1.907	0.058
Spring 1st	NEE02	43.400	8.152	5.324	0.000
Spring 1st	NEE04	-23.917	23.725	-1.008	0.314
Spring 1st	NEE05	-5.707	8.536	-0.669	0.504
Spring 1st	NEE11	-2.715	9.639	-0.282	0.778
Spring 1st	NEE13	7.194	10.165	0.708	0.480
Spring 1st	NEE15	2.008	10.356	0.194	0.846
Spring 1st	NEE81	-16.563	12.970	-1.277	0.203
Spring 1st	NEE82	14.498	13.167	1.101	0.272
Spring 1st	NEE83	-10.062	9.281	-1.084	0.279
Spring 1st	EDAVG	2.385	0.845	2.821	0.005
Spring 1st	FEMALE	9.045	4.589	1.971	0.050
Growth Rate 1-3	BASE	66.892	5.996	11.155	0.000
Growth Rate 1-3	PGMIS	3.169	6.339	0.500	0.618
Growth Rate 1-3	PONE13	-16.061	14.127	-1.137	0.257
Growth Rate 1-3	NIS01	-9.416	5.733	-1.642	0.102
Growth Rate 1-3	NIS10	-11.550	9.608	-1.202	0.231
Growth Rate 1-3	NIS7A	14.733	12.309	1.197	0.233
Growth Rate 1-3	NIS7B	8.936	8.733	1.023	0.307
Growth Rate 1-3	NIS7D	-14.357	11.865	-1.210	0.228
Growth Rate 1-3	NEE02	-9.072	5.992	-1.514	0.131
Growth Rate 1-3	NEE04	16.505	17.898	0.922	0.357
Growth Rate 1-3	NEE05	7.939	6.169	1.287	0.199
Growth Rate 1-3	NEE11	-19.939	6.378	-3.126	0.002
Growth Rate 1-3	NEE13	-2.745	7.133	-0.385	0.701
Growth Rate 1-3	NEE15	-25.504	6.550	-3.894	0.000
Growth Rate 1-3	NEE81	38.064	13.237	2.876	0.004
Growth Rate 1-3	NEE82	8.996	12.925	0.696	0.487
Growth Rate 1-3	NEE83	5.706	8.957	0.637	0.525
Growth Rate 1-3	FEMALE	1.145	3.458	0.331	0.741
Curvature 1-3	BASE	-17.471	5.120	-3.412	0.001
Curvature 1-3	PGMIS	-6.705	6.365	-1.053	0.293
Curvature 1-3	PONE13	21.736	14.371	1.513	0.132
Year-to-Year Deviations	ABSENT	-0.077	0.225	-0.340	0.734

Table 189

One-Program Schools 1-3 Analysis:
Two-Level Model for Reading (Pretest as Predictor)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	120.016	22.506	5.333	0.000
Spring 1st	PGMIS	1.599	6.320	0.253	0.800
Spring 1st	PONE13	25.038	19.595	1.278	0.203
Spring 1st	NIS01	0.224	6.360	0.035	0.972
Spring 1st	NIS10	28.725	12.965	2.216	0.028
Spring 1st	NIS7A	23.284	11.294	2.062	0.040
Spring 1st	NIS7B	-11.028	8.493	-1.299	0.195
Spring 1st	NIS7D	-21.202	11.167	-1.899	0.059
Spring 1st	NEE02	33.940	7.992	4.247	0.000
Spring 1st	NEE04	-29.839	22.570	-1.322	0.187
Spring 1st	NEE05	-6.427	8.105	-0.793	0.429
Spring 1st	NEE11	1.569	9.180	0.171	0.864
Spring 1st	NEE13	12.766	9.702	1.316	0.190
Spring 1st	NEE15	7.288	9.878	0.738	0.461
Spring 1st	NEE81	-10.393	12.375	-0.840	0.402
Spring 1st	NEE82	10.198	12.530	0.814	0.417
Spring 1st	NEE83	-14.327	8.861	-1.617	0.107
Spring 1st	EDAVG	2.454	0.800	3.069	0.002
Spring 1st	FEMALE	8.199	4.359	1.881	0.061
Spring 1st	ANALTOBE	0.241	0.051	4.737	0.000
Growth Rate 1-3	BASE	66.713	6.061	11.006	0.000
Growth Rate 1-3	PGMIS	3.217	6.381	0.504	0.615
Growth Rate 1-3	PONE13	-15.842	14.299	-1.108	0.269
Growth Rate 1-3	NIS01	-9.459	5.776	-1.638	0.103
Growth Rate 1-3	NIS10	-11.437	9.714	-1.177	0.240
Growth Rate 1-3	NIS7A	14.615	12.346	1.184	0.238
Growth Rate 1-3	NIS7B	8.877	8.762	1.013	0.312
Growth Rate 1-3	NIS7D	-14.214	11.908	-1.194	0.234
Growth Rate 1-3	NEE02	-9.015	6.058	-1.488	0.138
Growth Rate 1-3	NEE04	16.648	18.113	0.919	0.359
Growth Rate 1-3	NEE05	7.855	6.240	1.259	0.209
Growth Rate 1-3	NEE11	-19.962	6.454	-3.093	0.002
Growth Rate 1-3	NEE13	-2.763	7.213	-0.383	0.702
Growth Rate 1-3	NEE15	-25.422	6.631	-3.834	0.000
Growth Rate 1-3	NEE81	38.027	13.293	2.861	0.005
Growth Rate 1-3	NEE82	8.912	12.982	0.687	0.493
Growth Rate 1-3	NEE83	5.733	8.995	0.637	0.525
Growth Rate 1-3	FEMALE	1.263	3.500	0.361	0.719
Curvature 1-3	BASE	-17.654	5.121	-3.448	0.001
Curvature 1-3	PGMIS	-6.626	6.369	-1.040	0.299
Curvature 1-3	PONE13	22.038	14.381	1.532	0.127
Year-to-Year Deviations	ABSENT	-0.048	0.220	-0.219	0.827

Table 190

One-Program Schools 1-3 Analysis:
Two-Level Model for Reading (With Propensity for Pretest Students)

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	207.959	11.520	18.051	0.000
Spring 1st	PGMIS	-2.183	6.276	-0.348	0.728
Spring 1st	PONE13PN	24.491	17.559	1.395	0.164
Spring 1st	NIS01	2.588	6.490	0.399	0.690
Spring 1st	NIS10	23.308	12.305	1.894	0.059
Spring 1st	NIS7A	26.148	11.377	2.298	0.022
Spring 1st	NIS7B	-8.689	8.451	-1.028	0.305
Spring 1st	NIS7D	-22.036	11.062	-1.992	0.048
Spring 1st	NEE02	42.530	7.973	5.335	0.000
Spring 1st	NEE04	-24.506	23.123	-1.060	0.290
Spring 1st	NEE05	-6.270	8.316	-0.754	0.452
Spring 1st	NEE11	-3.097	8.787	-0.352	0.725
Spring 1st	NEE13	6.532	9.438	0.692	0.490
Spring 1st	NEE15	1.499	9.537	0.157	0.875
Spring 1st	NEE81	-15.395	12.290	-1.253	0.212
Spring 1st	NEE82	15.767	12.420	1.270	0.206
Spring 1st	NEE83	-9.097	8.606	-1.057	0.292
Spring 1st	EDAVG	2.373	0.790	3.005	0.003
Spring 1st	FEMALE	9.040	4.468	2.023	0.044
Growth Rate 1-3	BASE	64.886	5.524	11.746	0.000
Growth Rate 1-3	PGMIS	2.652	6.333	0.419	0.676
Growth Rate 1-3	PONE13PN	-13.318	13.133	-1.014	0.312
Growth Rate 1-3	NI[illegible]	-8.408	5.749	-1.463	0.145
Growth Rate 1-3	NIS[illegible]	-9.202	9.303	-0.989	0.324
Growth Rate 1-3	NIS7A	13.706	12.288	1.115	0.266
Growth Rate 1-3	NIS7B	7.927	8.696	0.912	0.363
Growth Rate 1-3	NIS7D	-16.537	11.704	-1.413	0.159
Growth Rate 1-3	NEE02	-8.422	6.090	-1.383	0.168
Growth Rate 1-3	NEE04	18.137	18.212	0.996	0.320
Growth Rate 1-3	NEE05	8.619	6.227	1.384	0.168
Growth Rate 1-3	NEE11	-18.405	6.208	-2.965	0.003
Growth Rate 1-3	NEE13	-1.151	6.996	-0.165	0.869
Growth Rate 1-3	NEE15	-23.915	6.446	-3.710	0.000
Growth Rate 1-3	NEE81	35.385	13.102	2.701	0.007
Growth Rate 1-3	NEE82	6.327	12.794	0.495	0.621
Growth Rate 1-3	NEE83	3.366	8.781	0.383	0.702
Growth Rate 1-3	FEMALE	1.077	3.519	0.306	0.760
Curvature 1-3	BASE	-16.076	4.912	-3.273	0.001
Curvature 1-3	PGMIS	-5.894	6.360	-0.927	0.355
Curvature 1-3	PONE13PN	16.850	13.991	1.204	0.230
Year-to-Year Deviations	ABSENT	-0.068	0.223	-0.305	0.761

Table 191

One-Program Schools 1-3 Analysis:
Two-Level Model for Reading
(With Pretest as Predictor of Propensity and Initial Status)

Predicted	Predictor	Parameter	Std Err	t-Stat	p value
Spring 1st	BASE	113.423	27.357	4.146	0.[illegible]0
Spring 1st	PGMIS	1.418	6.284	0.226	0.[illegible]
Spring 1st	PONE13PP	23.010	16.996	1.354	0.1[illegible]
Spring 1st	NIS01	-1.072	6.494	-0.165	0.86[illegible]
Spring 1st	NIS10	27.135	12.310	2.204	0.028
Spring 1st	NIS7A	23.848	11.317	2.107	0.036
Spring 1st	NIS7B	-9.917	8.342	-1.189	0.236
Spring 1st	NIS7D	-20.243	10.895	-1.858	0.064
Spring 1st	NEE02	33.605	8.186	4.105	0.000
Spring 1st	NEE04	-29.535	22.868	-1.292	0.198
Spring 1st	NEE05	-7.523	8.298	-0.907	0.366
Spring 1st	NEE11	0.456	8.835	0.052	0.959
Spring 1st	NEE13	11.665	9.521	1.225	0.222
Spring 1st	NEE15	5.873	9.521	0.617	0.538
Spring 1st	NEE81	-10.616	12.342	-0.860	0.391
Spring 1st	NEE82	12.760	12.222	1.044	0.298
Spring 1st	NEE83	-12.184	8.454	-1.441	0.151
Spring 1st	EDAVG	2.333	0.753	3.098	0.002
Spring 1st	FEMALE	8.260	4.413	1.872	0.062
Spring 1st	ANALTOBE	0.267	0.060	4.458	0.000
Growth Rate 1-3	BASE	67.140	5.019	13.377	0.000
Growth Rate 1-3	PGMIS	4.827	6.317	0.764	0.446
Growth Rate 1-3	PONE13PP	-20.860	11.494	-1.815	0.071
Growth Rate 1-3	NIS01	-8.222	5.678	-1.448	0.149
Growth Rate 1-3	NIS10	-11.792	8.856	-1.332	0.184
Growth Rate 1-3	NIS7A	14.612	12.242	1.194	0.234
Growth Rate 1-3	NIS7B	8.161	8.653	0.943	0.347
Growth Rate 1-3	NIS7D	-13.384	11.578	-1.156	0.249
Growth Rate 1-3	NEE02	-8.894	5.890	-1.510	0.132
Growth Rate 1-3	NEE04	15.207	17.874	0.851	0.396
Growth Rate 1-3	NEE05	9.370	6.132	1.528	0.128
Growth Rate 1-3	NEE11	-19.996	5.971	-3.349	0.001
Growth Rate 1-3	NEE13	-2.382	6.725	-0.354	0.724
Growth Rate 1-3	NEE15	-25.009	6.193	-4.038	0.000
Growth Rate 1-3	NEE81	41.062	13.264	3.096	0.002
Growth Rate 1-3	NEE82	6.661	12.626	0.528	0.598
Growth Rate 1-3	NEE83	3.845	8.576	0.448	0.654
Growth Rate 1-3	FEMALE	1.358	3.449	0.394	0.694
Curvature 1-3	BASE	-17.311	4.631	-3.738	0.000
Curvature 1-3	PGMIS	-7.546	6.482	-1.164	0.246
Curvature 1-3	PONE13PP	22.426	13.156	1.705	0.090
Year-to-Year Deviations	ABSENT	-0.049	0.220	-0.220	0.826

Summary

The immersion strategy students in one-program schools who have pretest scores available show greater mathematics and language growth rates than the early-exit students with pretest scores available, but the reading growth curves are essentially the same for the two programs. More importantly, the conclusions drawn from the HLM results do not change when pretest is included in the analysis from when pretest is excluded from the analysis. (e.g., compare p-values for PGMIS from Table 190 and Table 191). Establishing this insensitivity was the main purpose of the 1-3 pretest analyses; the result tends to support the appropriateness of the basic model, which does not include the pretest.

Results of the Three-Level 1-3 Analyses

What were the results of the three-level 1-3 analysis for the mathematics subtest?

The two-level HLM program does not permit the treatment of both student-level and school-level parameters as random effects. For the K-1 analyses, schools were treated as random (i.e., an individual growth curve for each school was estimated) but students did not need to be treated as random because each student had only one test score, the spring first grade score (i.e., there is no need to estimate a growth curve for individual students having only one data point). For the two-level 1-3 analyses, the individual growth curves fitted to each individual had both an intercept (status at first grade) and a slope (growth rate from first grade to third grade) and these parameters were treated as random. Thus schools had to be treated as fixed effects. This can produce possibly misleading results, as described above. Accordingly, the three-level program was used to allow for random school effects.

For the three-level analyses of one-program schools, as for the corresponding two-level analyses, each student had an individual growth curve estimated. The individual intercepts and slopes were then predicted

ERIC

using the propensity score and background variables and a school effect. In the simplest models, the school effects were then predicted using a single binary variable for the difference between the immersion strategy and early-exit programs. In the models labeled "district main effects," the school effects were predicted using the program variable and district variables. In the models labeled "district main effects and interactions," the school effects were predicted using the program variable, the district variables, and variables indicating district-program interactions.

Table 192 summarizes the results of the simplest three-level model for the mathematics subtest. The first column, labeled "Time Within Student" shows the aspect of the student's individual growth curve that is being estimated. The "Spring 1st" line indicates the first line of several that comprise the model for initial status (achievement at spring of first grade). The "Growth 1-3" line indicates the first of several lines that comprise the model for the individual student slope (growth rate) from spring of first grade to spring of third grade. The "Curvature" line marks the part of the model used to predict the amount of curvature. As for the two-level HLM models, individual differences in curvature could not be accommodated because there were only two or three data points per student; therefore, the curvature is a fixed effect.

Table 192

One-Program Schools 1-3 Analysis for Mathematics
Without District

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	247.883	6.980	35.516	0.000
	BASE	PGMIS	2.061	5.849	0.352	0.730
	PONE13	BASE	0.207	8.442	0.025	0.980
	AVABS13	BASE	-0.619	0.205	-3.025	0.003
	EDAVG	BASE	1.652	0.398	4.156	0.000
	FEMALE	BASE	-4.821	2.554	-1.887	0.060
Growth 1-3	BASE	BASE	45.717	3.954	11.561	0.000
	BASE	PGMIS	-3.058	4.937	-0.619	0.536
	PONE13	BASE	-1.396	5.792	-0.241	0.810
	FEMALE	BASE	5.879	1.723	3.413	0.001
Curvature	BASE	BASE	-6.823	2.706	-2.521	0.012
	BASE	PGMIS	-3.818	2.624	-1.455	0.146
	PONE13	BASE	12.481	5.079	2.457	0.014

The next column, labeled "Student Within School," indicates the predictors of between-student variation within schools. The predictors include PONE13 (the propensity score), EDAVG (parents' average years of education), and FEMALE (a binary variable that is 1 for girls and 0 for boys). Other than the school variables, these are the same predictors as in the two-level model (see Table 174). In addition, AVABS13 (the average number of absences in grades 1 through 3) is included as a predictor of initial status instead of having ABSENT as a separate within-student variable. As explained in Chapter II, the inclusion of AVABS13 as a predictor of initial status is similar to including ABSENT as a within-student variable. With the three-level program, the missing data for ABSENT meant it could not be included in the analysis.

The "BASE" for students within each school represents the school effects. The BASE for the Spring 1st lines represents the school effect on the initial status, the BASE for the Growth 1-3 lines represents the

ERIC

school effect on the growth rate, and the BASE for the Curvature lines represents the school effect on the curvature. As described above, the amount of school variation in curvature was assessed using the two-level program. The school variation for curvature was found to be only marginally statistically significant and much smaller than the school variability in initial status and growth rate. Accordingly, the three-level model was run with school differences allowed only for the initial status and growth rate; curvature was treated as fixed across the schools.

The third column, labeled "Predictor of School," indicates the variables used to predict school differences. The only predictor other than an overall constant in this relatively simple model is PGMIS, the indicator of program membership for the school. The remaining columns, giving the parameter estimates and related values, are interpreted in the usual way.

The estimated program difference in initial status (after adjusting for propensity) is given by the parameter value associated with PGMIS in the "Spring 1st" group of lines. It shows a very small immersion strategy advantage for the mathematics subtest, only 2.061 points, which is not significant ($p = .730$). The estimated program difference in growth rate is given by the coefficient of PGMIS in the next section of the table: it shows a nonsignificant advantage for early-exit schools (difference = -3.058, $p = .536$). The estimated program difference in curvature is also nonsignificant (-3.818, $p = .146$).

Table 192 can be compared with the corresponding table for the two-level model, Table 174. The estimated program differences are fairly similar, but the standard errors for the program differences in initial status and growth rate are noticeably larger for the three-level model (5.849 instead of 3.179, and 5.719 instead of 2.186, respectively). This reflects the fact that the three-level model correctly incorporates school variability in the estimated standard error. The standard error of the program effect for curvature is similar in the two-level and three-level models (in this case, both are 2.624) because curvature is not allowed to vary by school in either model.

Table 193 shows the three-level model with district main effects added as predictors of school variability. The district coding is the same as used in the K-1 models. District C is the omitted district and each of the DIST variables represents the difference between that district and district C. The hypothesis that there are no district differences in initial status or growth rate can be tested using a likelihood ratio chi-square test; the district differences are marginally significant at the .05 level (chi-square = 16.8 with 8 d.f., p = .032).

Table 193

One-Program Schools 1-3 Analysis for Mathematics
With District Main Effects

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	258.879	10.349	25.013	0.000
	BASE	PGMIS	2.153	5.719	0.376	0.713
	BASE	DISTA	0.580	9.484	0.061	0.952
	BASE	DISTB	-14.171	10.054	-1.409	0.182
	BASE	DISTF	18.660	15.398	1.212	0.247
	BASE	DISTHI	3.228	9.130	0.354	0.729
	PONE13	BASE	-16.360	11.240	-1.456	0.146
	AVABS13	BASE	-0.611	0.206	-2.967	0.003
	EDAVG	BASE	1.261	0.434	2.906	0.004
	FEMALE	BASE	-4.753	2.549	-1.865	0.063
Growth 1-3	BASE	BASE	37.490	6.669	5.621	0.000
	BASE	PGMIS	-3.970	4.276	-0.928	0.354
	BASE	DISTA	16.349	7.080	2.309	0.021
	BASE	DISTB	2.155	7.440	0.290	0.772
	BASE	DISTF	9.485	10.900	0.870	0.385
	BASE	DISTHI	14.579	6.660	2.189	0.029
	PONE13	BASE	-5.854	6.744	-0.868	0.386
	FEMALE	BASE	5.958	1.717	3.469	0.001
Curvature	BASE	BASE	-6.666	2.713	-2.457	0.014
	BASE	PGMIS	-3.894	2.624	-1.484	0.139
	PONE13	BASE	12.279	5.089	2.413	0.016

Districts A and H/I have the highest mathematics growth rates and districts B and C the lowest. District F has growth rates that are relatively high but not significantly different from any of the other districts. The estimated district differences in initial status are not large compared to their standard errors, confirming the finding from the K-1 analyses that districts do not differ significantly in spring first grade mathematics achievement. To the extent there are differences, district B is low and district F is high in spring of first grade.

The three-level program was used to estimated a model for the mathematics subtest that included district main effects and district-program interactions. The district-program interactions were not significant (chi-square = 4.9 with 6 d.f., p = .557) so that model is not presented.

Summary

The three-level HLM model confirms the finding of no statistically significant differences between the programs on the mathematics subtest for the one-program schools after adjusting for propensity. The district differences are marginally significant. Districts A and H/I have relatively high mathematics growth rates and districts B and C the lowest. The mathematics growth rate in district F is fairly high but not significantly different from the growth rates in districts B and C. The estimated initial status (spring of first grade) is low in district B and high in district F, but the district differences in initial status are small compared with their standard errors and therefore are not statistically significant.

What were the results of the three-level 1-3 analysis for the language subtest?

Table 194 presents the results of the three-level 1-3 analysis for the language subtest, with school random and no district effects included. The program difference in initial status and growth rate both show a small immersion strategy advantage (difference = 3.120 in initial status, difference = 5.997 in growth rate) but neither difference is statistically significant (p = .747 and p = .335, respectively). These findings confirm the results from the two-level model (see Table 175). Also confirmed in the three-level model is the statistically significant program difference in curvature (difference = -14.487, p = .001). The larger negative curvature of the immersion strategy program together with the nonsignificant difference in initial status and growth rate confirms the two-level model finding that IS schools show a temporary increase in language achievement. The estimated program differences for first grade through third grade are -1.709, 18.775, and 10.285.

Table 194

One-Program Schools 1-3 Analysis for Language Without District

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	233.930	11.208	20.872	0.000
	BASE	PGMIS	3.120	9.505	0.328	0.747
	PONE13	BASE	25.115	13.108	1.916	0.056
	AVABS13	BASE	-0.524	0.330	-1.586	0.114
	EDAVG	BASE	1.990	0.646	3.080	0.002
	FEMALE	BASE	3.239	4.017	0.806	0.421
Growth 1-3	BASE	BASE	61.391	5.453	11.258	0.000
	BASE	PGMIS	5.997	6.215	0.965	0.335
	PONE13	BASE	-14.447	8.729	-1.655	0.099
	FEMALE	BASE	14.591	3.023	4.827	0.000
Curvature	BASE	BASE	-15.542	4.380	-3.549	0.000
	BASE	PGMIS	-14.487	4.260	-3.400	0.001
	PONE13	BASE	25.616	8.217	3.117	0.002

Table 195 shows the three-level model for the language subtest with district main effects included. The district effects are significant (chi-square = 23.0 with 8 d.f., p = .003). The program effects are little changed; the estimated program differences for the three grades are now 0.892, 21.466, and 11.584. Districts F and H/I have relatively high growth rates and districts B and C relatively low growth rates. District A has a language growth rate in between those of the two pairs of districts, but closer to the higher pair (F and H/I). Districts A and F have high initial status and district C low initial status, confirming those results from the K-1 analysis of one-program schools. District H/I has nearly as high an initial status as districts A and F, which was also found in the K-1 analysis. However, in this 1-3 analysis district B is nearly as high in initial status as district H/I, whereas in the K-1 analysis district B was nearly as low as district C. The difference between district B and district C is not statistically significant in this 1-3 analysis (p = .167), so the findings are not in conflict.

Table 195

One-Program Schools 1-3 Analysis for Language
With District Main Effects

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	206.434	14.624	14.116	0.000
	BASE	PGMIS	5.968	7.467	0.799	0.439
	BASE	DISTA	44.938	12.158	3.696	0.003
	BASE	DISTB	19.366	13.233	1.463	0.167
	BASE	DISTF	37.472	20.870	1.795	0.096
	BASE	DISTHI	22.296	12.088	1.845	0.088
	PONE13	BASE	25.433	17.713	1.436	0.152
	AVABS13	BASE	-0.481	0.332	-1.450	0.148
	EDAVG	BASE	2.006	0.701	2.862	0.004
	FEMALE	BASE	3.320	4.003	0.829	0.408
Growth 1-3	BASE	BASE	59.471	8.180	7.270	0.000
	BASE	PGMIS	5.346	4.954	1.079	0.281
	BASE	DISTA	14.312	7.751	1.846	0.066
	BASE	DISTB	-5.920	8.693	-0.681	0.496
	BASE	DISTF	26.488	13.137	2.016	0.044
	BASE	DISTHI	22.754	7.702	2.954	0.003
	PONE13	BASE	-36.336	11.516	-3.155	0.002
	FEMALE	BASE	14.689	2.996	4.902	0.000
Curvature	BASE	BASE	-14.554	4.409	-3.301	0.001
	BASE	PGMIS	-15.228	4.269	-3.567	0.000
	PONE13	BASE	24.764	8.264	2.996	0.003

The model including district-program interactions showed statistically significant interactions (chi-square = 10.4 with 6 d.f., p = .109), so that model is not presented here.

<u>Summary</u>

The three-level 1-3 analysis for the language subtest confirms the earlier findings. After adjusting for propensity, no significant program differences in initial status or average growth were detected, but the immersion strategy program has greater curvature than the early-exit program. Overall district differences on the language subtest are statistically significant. The differences in initial status are similar to those found in the K-1 analyses. The language growth rate in districts B and C are lower than the growth rates in districts F and H/I. District A has growth rates in between the two pairs of districts, but closer to districts F and H/I.

<u>What were the results of the three-level 1-3 analysis for the reading subtest?</u>

The results of the three-level 1-3 analysis without district effects are summarized in Table 196. There are no significant program differences in initial status, growth rate, or curvature. The estimated difference between the programs in first grade is -6.549, in second grade it is 1.620, and in third grade it is 1.697. These results are similar to the findings from the two-level analysis (see Table 176).

Table 196

One-Program Schools 1-3 Analysis for Reading
Without District

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	228.476	9.061	25.216	0.000
	BASE	PGMIS	-5.200	7.225	-0.720	0.481
	PONE13	BASE	1.020	10.725	0.095	0.924
	AVABS13	BASE	-0.389	0.281	-1.387	0.166
	EDAVG	BASE	1.831	0.543	3.374	0.001
	FEMALE	BASE	3.955	3.429	1.153	0.250
Growth 1-3	BASE	BASE	54.757	4.366	12.541	0.000
	BASE	PGMIS	4.123	4.833	0.853	0.394
	PONE13	BASE	-3.110	7.149	-0.435	0.664
	FEMALE	BASE	4.544	2.549	1.782	0.075
Curvature	BASE	BASE	-18.340	3.907	-4.694	0.000
	BASE	PGMIS	-4.046	3.789	-1.068	0.286
	PONE13	BASE	24.581	7.338	3.350	0.001

Table 197 shows the results for the reading subtest with district main effects included. The estimated program differences are little changed from the model without district main effects. The estimated program differences for first through third grade are now -5.678, 2.529, and 2.284. The test for overall district effects was not significant, indicating the districts have comparable initial status and growth rates (chi-square = 9.8 with 8 d.f., p = .279).

Table 197

One-Program Schools 1-3 Analysis for Reading
With District Main Effects

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	211.545	12.694	16.665	0.000
	BASE	PGMIS	-4.269	6.561	-0.651	0.526
	BASE	DISTA	27.378	10.719	2.554	0.024
	BASE	DISTB	14.093	11.609	1.214	0.246
	BASE	DISTF	14.969	18.221	0.822	0.426
	BASE	DISTHI	14.022	10.599	1.323	0.209
	PONE13	BASE	2.581	15.130	0.171	0.864
	AVABS13	BASE	-0.359	0.283	-1.266	0.206
	EDAVG	BASE	1.813	0.598	3.032	0.003
	FEMALE	BASE	3.938	3.419	1.152	0.250
Growth 1-3	BASE	BASE	57.605	7.630	7.550	0.000
	BASE	PGMIS	3.981	4.709	0.845	0.399
	BASE	DISTA	4.584	7.497	0.611	0.542
	BASE	DISTB	-7.233	8.228	-0.879	0.380
	BASE	DISTF	12.693	12.306	1.031	0.303
	BASE	DISTHI	7.618	7.317	1.041	0.298
	PONE13	BASE	-16.016	9.943	-1.611	0.108
	FEMALE	BASE	4.660	2.572	1.812	0.071
Curvature	BASE	BASE	-17.733	3.901	-4.546	0.000
	BASE	PGMIS	-4.226	3.774	-1.120	0.263
	PONE13	BASE	24.017	7.314	3.284	0.001

However, the test for district-program interactions was significant at the .05 level (chi-square = 15.3 with 6 d.f., p = .018). The combined test for district main effects and district-program interactions was also significant (chi-square = 25.1 with 14 d.f., p = .034). Thus the null hypothesis that the district differences are the same in all the districts must be rejected at the .05 level (although not at the .01 level). The model with district-program interaction dummy variables is complicated. The coefficients of the propensity score and background variables in the interaction model are essentially the same as in the district main effects model. Therefore, the results of the interaction model that are of primary interest are the estimated initial status and growth rate in each of the district-program combinations. These are given in Table 198.

Table 198

One-Program Schools 1-3 Analysis for Reading
Estimated District-Program Effects Relative to Early-Exit in District C

		Early-Exit	Immersion Strategy	Difference IS - EE
Spring 1st	District A	29.391	7.077	-22.314
	District B	4.973	24.014	19.041
	District C	*	-14.824	-14.824
	District F	10.982	—	—
	District H/I	6.044	8.129	2.085
Growth 1-3	District A	16.878	20.166	-18.177
	District B	4.979	12.353	31.350
	District C	*	30.947	44.631
	District F	24.385	—	—
	District H/I	25.389	-5.326	-11.006

* The early-exit school group in district C serves as the reference category; all other values are differences from this school group.

The district-program interaction appears to be due in large part to much higher reading growth rate for the IS school group in district C than for the EE school group in that district. The growth rates in the other districts do not differ greatly between the IS and EE schools. In light of the marginal significance of the district-program interactions, the estimates in Table 198 should not be overinterpreted. The estimated initial status is especially unstable.

The presence of district-program interactions considerably complicates the interpretation of the results. An alternative to the full district main effects and district-by-program interaction model is a model that allows for district differences within the EE program but not within the IS program. This model is a plausible alternative in part because the amount of instructional English varies substantially among the districts for the EE program but scarcely varies at all among the districts for the IS program.

For the reading subtest, the alternative model of district differences for early-exit schools but not immersion strategy schools was fit to the data using the three-level HLM program. District differences on the reading subtest were significant among the EE programs (chi-square = 16.2 with 8 d.f., p = .040). Recall that the overall district differences were not significant (p = .279). Furthermore, the test that IS programs also had district differences was not significant (chi-square = 8.9 with 6 d.f., p = .179). This alternative model is preferable to the full district-by-program interaction model because it is more parsimonious and fits the data nearly as well.

Table 199 gives the results of estimating this alternative model of EE district differences but no IS district differences. The early-exit school group in district C serves as the reference. The PGMIS parameter estimate gives the amount by which the immersion strategy schools differ from the EE district C school group. Recall that district F has no immersion strategy school, so it does not contribute to the estimate of the PGMIS parameter. The EE-A, EE-B, EE-F, and EE-HI parameter estimates give the difference between the EE school group in district C and the EE program in each district respectively.

Table 199

One-Program Schools 1-3 Analysis for Reading
With District Effects for Early-Exit Only

		Early-Exit	Immersion Strategy	Difference IS - EE
Spring 1st	District A	29.391	7.077	-22.314
	District B	4.973	24.014	19.041
	District C	*	-14.824	-14.824
	District F	10.982	—	—
	District H/I	6.044	8.129	2.085
Growth 1-3	District A	16.878	20.166	-18.177
	District B	4.979	12.353	31.350
	District C	*	30.947	44.631
	District F	24.385	—	—
	District H/I	25.389	-5.326	-11.006

* The early-exit school group in district C serves as the reference category; all other values are differences from this school group.

Under this alternative model, the growth rates for the IS programs as a group and the EE programs in districts A, F, and H/I are all about equal and substantially larger than the growth rates in the EE programs in districts B and C. The estimated initial status is similar across all the districts and programs except that the EE program in district A is relatively high. The difference between the EE program in district A and the lowest initial status, for the EE program in the reference district C, falls just short of statistical significance (p = .055).

It is noteworthy that the districts with highest early-exit reading growth rate, districts A, F, and H/I, are the three districts with the highest use of English for instruction in their early-exit programs (see Chapter I). The two districts with relatively low use of English in their early-exit program, districts B and C, both show relatively low reading growth rates. An evaluation of this suggestive pattern goes beyond nominal program analyses.

Summary

The program differences on the reading subtest are small after adjusting for propensity, but there is some evidence that the program effect varies by district. A full district-by-program interaction model, fitting a separate initial status and growth rate to each district-program combination, was estimated for the reading subtest. For that model, the reading growth rate for the immersion strategy school group in district C is found to be much higher than for the early-exit school group in that district. In the other three districts with both programs represented among the one-program schools, the program difference in growth rate is much smaller. An alternative, more parsimonious, model was also fit to the reading subtest. In this alternative model, early-exit schools differ by district but immersion strategy schools do not. The hypothesis that the growth curves are the same for all four districts with IS schools cannot be rejected. In this model, all of the district-program combinations have approximately the same growth rate except for the early-exit schools in districts B and C, which have lower growth rates. The early-exit programs in districts B and C have the lowest use of English for instruction among these districts and programs. Both models indicate that all of the districts and programs have approximately the same initial status.

Does the alternative model used for the reading subtest also improve the fit for the mathematics and language subtests?

For the mathematics and language subtests, district-by-program interactions were not statistically significant. A model of main effects for program and main effects for district provided an adequate fit to the schools. Because of the significant district-by-program interactions for the reading subtest, an alternative model of EE district differences but no IS district differences was fit to the reading data and found to provide an adequate fit with a simpler model than the full interaction model. For comparison purposes, it is useful to check the fit of the alternative model for the mathematics and language subtests.

For both the mathematics and the language subtests, the original model with district main effects and the alternative model with district differences only for early-exit schools fit the data about equally well. Table 200 and Table 201 summarize the results of the model-testing for the mathematics and language subtests, respectively. For comparison, Table 202 gives the corresponding results for the reading subtest.

Table 200

One-Program Schools 1-3 Analysis for Mathematics
Tests of District and District-Program Effects

	Likelihood Ratio Chi-Square	Degrees of Freedom	p-Value
District Main Effects vs. No District Effects	16.8	8	0.032
District-Program Interactions vs. District Main Effects	4.9	6	0.557
District Main Effects and District-Program Interactions vs. No District Effects	21.7	14	0.085
EE District Effects vs. No District Effects	16.3	8	0.038
EE and IS District Effects vs. EE District Effects	5.4	6	0.494

Table 201

One-Program Schools 1-3 Analysis for Language
Tests of District and District-Program Effects

	Likelihood Ratio Chi-Square	Degrees of Freedom	p-Value
District Main Effects vs. No District Effects	23.0	8	0.003
District-Program Interactions vs. District Main Effects	10.4	6	0.109
District Main Effects and District-Program Interactions vs. No District Effects	33.4	14	0.003
EE District Effects vs. No District Effects	24.8	8	0.002
EE and IS District Effects vs. EE District Effects	8.6	6	0.197

Table 202

One-Program Schools 1-3 Analysis for Reading
Tests of District and District-Program Effects

	Likelihood Ratio Chi-Square	Degrees of Freedom	p-Value
District Main Effects vs. No District Effects	9.8	8	0.279
District-Program Interactions vs. District Main Effects	15.3	6	0.018
District Main Effects and District-Program Interactions vs. No District Effects	25.1	14	0.034
EE District Effects vs. No District Effects	16.2	8	0.040
EE and IS District Effects vs. EE District Effects	8.9	6	0.179

Table 203 and Table 204 show the mathematics and language models estimated with district differences for the EE program but not the IS program. They may be compared with the district main effects models in Table 193 and Table 195, respectively. The results for the mathematics subtest indicate fairly small differences among the early-exit districts, which is consistent with the marginal statistical significance of the district effects. For the language subtest, the pattern of results with the alternative model is very similar to the pattern for the reading subtest. The growth rates for the early-exit schools in districts A, F, and H/I are very similar to the growth rate estimated for the immersion strategy schools, while the growth rates for districts B and C are lower. The initial status results indicate the early-exit programs in districts B, C, and H/I start somewhat lower than the EE programs in districts A and F. The spring first grade status for the IS programs, which under this model is the same for all four districts with IS-only schools (A, B, C,

ERIC

and H/I), is comparable to the status for the EE programs in districts A and F.

Table 203

One-Program Schools 1-3 Analysis for Mathematics
With District Effects for Early-Exit Only

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	249.878	11.685	21.385	0.000
	BASE	PGMIS	7.273	11.139	0.653	0.525
	BASE	EE-A	10.222	12.009	0.851	0.410
	BASE	EE-B	-9.707	12.171	-0.798	0.439
	BASE	EE-F	20.661	16.040	1.288	0.220
	BASE	EE-HI	6.325	11.955	0.529	0.606
	PONE13	BASE	-10.711	9.956	-1.076	0.283
	AVABS13	BASE	-0.553	0.206	-2.685	0.008
	EDAVG	BASE	1.424	0.413	3.451	0.001
	FEMALE	BASE	-4.692	2.547	-1.842	0.066
Growth 1-3	BASE	BASE	38.694	9.160	4.224	0.000
	BASE	PGMIS	5.163	9.460	0.546	0.585
	BASE	EE-A	14.519	10.126	1.434	0.152
	BASE	EE-B	-0.658	10.278	-0.064	0.949
	BASE	EE-F	6.022	12.910	0.466	0.641
	BASE	EE-HI	12.184	10.011	1.217	0.224
	PONE13	BASE	-3.474	6.403	-0.543	0.587
	FEMALE	BASE	5.929	1.715	3.457	0.001
Curvature	BASE	BASE	-6.648	2.717	-2.447	0.015
	BASE	PGMIS	-3.924	2.630	-1.492	0.136
	PONE13	BASE	12.259	5.095	2.406	0.017

Table 204

One-Program Schools 1-3 Analysis for Language
With District Effects for Early-Exit Only

Time Within Student	Student Within School	Predictor of School	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	BASE	206.939	14.793	13.989	0.000
	BASE	PGMIS	30.079	13.179	2.282	0.040
	BASE	EE-A	55.288	14.200	3.894	0.002
	BASE	EE-B	14.004	14.535	0.963	0.353
	BASE	EE-F	39.497	19.951	1.980	0.069
	BASE	EE-HI	17.510	14.404	1.216	0.246
	PONE13	BASE	22.658	14.975	1.513	0.131
	AVABS13	BASE	-0.430	0.332	-1.293	0.197
	EDAVG	BASE	2.020	0.659	3.068	0.002
	FEMALE	BASE	3.355	4.013	0.836	0.404
Growth 1-3	BASE	BASE	48.071	10.326	4.655	0.000
	BASE	PGMIS	25.868	10.230	2.529	0.012
	BASE	EE-A	21.363	10.808	1.977	0.049
	BASE	EE-B	3.716	11.228	0.331	0.741
	BASE	EE-F	27.145	14.504	1.872	0.062
	BASE	EE-HI	31.625	10.993	2.877	0.004
	PONE13	BASE	-25.558	10.386	-2.461	0.014
	FEMALE	BASE	14.716	3.005	4.897	0.000
Curvature	BASE	BASE	-14.769	4.400	-3.356	0.001
	BASE	PGMIS	-15.451	4.266	-3.622	0.000
	PONE13	BASE	25.397	8.239	3.083	0.002

Summary

The original model of district main effects and the alternative model of district effects only for early-exit programs fit equally well to the mathematics and language subtests. The choice between these models cannot be made on statistical grounds. The pattern of growth rate results for the language subtest are similar to the pattern for the reading subtest. This pattern is suggestive of a relationship between the greater amounts of English and higher growth rates. The evaluation of this possible relationship goes beyond the scope of this report.

ERIC

Matching the K-1 and 1-3 Analyses

How well do the K-1 and 1-3 analyses for mathematics match in spring of first grade?

The K-1 and 1-3 analyses both produce estimated program differences for spring of first grade. The estimates from the K-1 analyses are based on analysis of covariance (ANCOVA) with school as a random effect. The estimates from the main set of 1-3 analyses are based on two-level hierarchical linear models fitted to individual growth curves, with school as a fixed effect. Some additional 1-3 analyses were performed using three-level hierarchical linear models, with school as a random effect. If the models provide reasonable fits to the data, all the analytic methods should produce similar program estimates for spring of first grade. Table 205 summarizes some of the estimated program differences for spring of first grade for the mathematics subtest.

Table 205

One-Program Schools K-1 and 1-3 Analysis:
Estimated Program Differences in Spring 1st for Mathematics

Description	Parameter	Std Err	t-Stat	p-Value
K-1 All: No District Effects	1.208	7.222	0.167	0.870
K-1 All: With District Main	0.813	7.640	0.106	0.918
K-1 Pre: No District Effects	0.626	8.449	0.074	0.942
K-1 Pre: With District Main	-0.865	8.330	-0.104	0.920
K-1 Pre: Mathematics Pretest	1.802	7.428	0.243	0.814
K-1 Pre: District & Pretest	0.453	7.953	0.057	0.956
1-3 All: Basic Two-Level	1.960	3.27*	0.599	0.550
1-3 All: School Curvature	2.318	3.2**		
1-3 Pre: Basic Two-Level	-3.274	4.6**		
1-3 Pre: Mathematics Pretest	-1.672	4.3**		
1-3 Pre: Propensity for Pre	-4.545	4.5**		
1-3 Pre: Pre Predicts Both	-1.920	4.2**		
1-3 Three-Level: No District	0.788	5.89*	0.134	0.895
1-3 Three-Level: District Main	0.855	5.77*	0.148	0.885
1-3 Three-Level: District EE Effects Only				
District A	-4.257	12.0**		
District B	15.672	12.2**		
District C	5.965	11.16*	0.534	0.602
District F	-14.696	16.0**		
District H/I	-0.360	12.0**		

* Estimated from t-statistic

** No t-statistic was calculated, so the standard error of the coefficient for the dummy variable as a predictor of initial status (intercept) is given.

The estimated program difference is given for six of the K-1 analyses, one from each of the six sets of four analyses. The analysis controlling for "key covariates" was selected as the representative of the set of four analyses. The estimated program difference at spring of first grade is also given for the six basic (two-level) 1-3 analyses and the three 1-3 analyses using the three-level HLM program; these estimated program differences at spring of first grade incorporate the curvature for the estimated program difference.

All of the K-1 analyses given indicate extremely small program differences in mathematics at spring of first grade. The basic 1-3 analyses show slightly larger differences but the largest difference is only about one standard error, which is not even close to statistical significance. Therefore, the fact that the standard errors for the basic analyses, which treat school as a fixed effect, are known to be underestimates of the true standard error does not change the conclusion that there are no program differences at spring of first grade for mathematics.

The 1-3 analyses using the three-level program, which correctly treat school as a random effect, also show nonsignificant differences between the programs. For the model with no district effects and for the model with district main effects, a single program difference is estimated and it is less than one ESS point, far from statistical significance. The alternative model with district effects for the EE program only produces a separate estimate of the program difference for each district. The estimated spring first grade mathematics achievement level for the IS program is the same for all districts, but the estimated spring first grade mathematics achievement level for the EE program varies by district (see Table 203). Therefore, the estimated difference between the IS program (estimated using all four districts with IS schools) and the EE program (estimated separately for each district) is different for each district. The difference for district F is presented in the table even though district F had no one-program immersion strategy schools; district F therefore did not contribute to the estimated first grade mathematics score for immersion strategy programs.

Because all of the 1-3 analyses allow the amount of curvature to vary by program, the estimated program difference in spring of first grade must be tested using the multivariate hypothesis testing capability of the HLM program. Only some of the possible hypothesis tests were performed. Where the hypothesis test was not performed, a rough approximation to the

ERIC

standard error of the estimated program difference can be obtained from the standard error of the dummy variable as a predictor of the initial status (intercept). For these analyses, that standard error tends to be a slight underestimate of the standard error of the predicted difference at spring of first grade. Only the largest estimated program difference, the 15.7 point EE advantage in district B, exceeds this rough standard error and only by a small margin. The difference would have to be at least twice the standard error to achieve statistical significance.

Summary

The K-1 and 1-3 estimates of program differences in spring of first grade for the mathematics subtest match well. The estimates are all small and far from statistical significance.

How well do the K-1 and 1-3 analyses for language match in spring of first grade?

The estimated program differences for the six "Key Covariates" K-1 analyses and the various 1-3 analyses for the language subtest are summarized in Table 206. The K-1 analyses indicate a modest but fairly consistent IS advantage in spring of first grade, but none of the estimates are close to statistical significance. The 1-3 analyses that produce a single overall estimate of the program difference for language in spring of first grade show small differences less than half the standard error (and therefore far from statistical significance).

ERIC

Table 206

One-Program Schools K-1 and 1-3 Analysis:
Estimated Program Differences in Spring 1st for Language

Description		Parameter	Std Err	t-Stat	p-Value
K-1 All: No District Effects		3.507	10.351	0.339	0.740
K-1 All: With District Main		7.978	9.188	0.868	0.408
K-1 Pre: No District Effects		8.404	12.444	0.675	0.515
K-1 Pre: With District Main		9.955	11.683	0.852	0.422
K-1 Pre: Pretest Total		9.103	10.524	0.865	0.410
K-1 Pre: District & Pretest		11.106	10.378	1.070	0.326
1-3 All: Basic Two-Level		-0.878	5.16*	-0.170	0.865
1-3 All: School Curvature		0.432	5.1**		
1-3 Pre: Basic Two-Level		-2.954	6.9**		
1-3 Pre: Pretest Total		1.338	6.6**		
1-3 Pre: Propensity for Pre		-3.125	6.7**		
1-3 Pre: Pre Predicts Both		1.594	6.6**		
1-3 Three-Level: No District		-1.709	9.58*	-0.178	0.861
1-3 Three-Level: District Main		0.892	7.55*	0.118	0.908
1-3 Three-Level: District EE Effects Only					
	District A	-30.359	14.2**		
	District B	10.925	14.5**		
	District C	24.929	13.22*	1.886	0.082
	District F	-14.568	20.0**		
	District H/I	7.419	14.4**		

* Estimated from t-statistic

** No t-statistic was calculated, so the standard error of the coefficient for the dummy variable as a predictor of initial status (intercept) is given.

For the alternative model with district differences included only for the EE programs, the estimates of program differences in the districts vary considerably. There is an estimated 25 point advantage for the IS programs compared with the EE program in district C, while there is an estimated 30 point advantage for the EE program in district A compared with the IS programs as a group. These differences approach statistical significance at the .05 level, since they are about twice the standard error. Undoubtedly the large difference between the EE programs in these two districts contributes greatly to the statistically significant district differences found for the language subtest. The other three districts have estimated spring first grade achievement levels for the EE schools that are well within one standard error of the estimated level for IS schools. These differences average out, however, so that the overall program difference (in the three-level models that estimate a single value for the program differences) are less than two points in either direction.

Summary

Both the K-1 and 1-3 analyses indicate no statistically significant overall program differences on the language subtest. Using the alternative model of district differences for EE schools but not for IS schools, substantial differences are observed that are near statistical significance at the .05 level. The EE program in district A exhibits higher language achievement at spring of first grade than the IS schools as a group. The opposite is true in district C: the EE program in district C exhibits lower first grade language subtest scores than the IS program schools.

How well do the K-1 and 1-3 analyses for reading match in spring of first grade?

The estimated program differences on the reading subtest for the six "Key Covariates" K-1 analyses and all the 1-3 analyses are given in Table 207. The K-1 analyses show small and nonsignificant differences in spring of first grade for the reading subtest. The 1-3 analyses on all students treating school as a fixed effect ("1-3 All: Basic Two-Level" and "1-3 All: School Curvature") show an early-exit advantage that approaches statistical significance. Because the standard error of the estimated program difference is an underestimate for the two-level model, this implies the difference is definitely not statistically significant. This is confirmed by the 1-3 analyses using the three-level program with and without district main effects ("1-3 Three-Level: No District" and "1-3

Three-Level: District Main"). In these two analyses, the early-exit advantage is similar to the program difference found in the analyses treating school as a fixed effect, but with the larger (and more accurate) standard error the difference is far from statistical significance.

ERIC

Table 207

One-Program Schools K-1 and 1-3 Analysis:
Estimated Program Differences in Spring 1st for Reading

Description		Parameter	Std Err	t-Stat	p-Value
K-1 All: No District Effects		-2.586	7.856	-0.329	0.747
K-1 All: With District Main		-2.019	8.177	-0.247	0.811
K-1 Pre: No District Effects		4.457	10.335	0.431	0.675
K-1 Pre: With District Main		4.885	11.392	0.429	0.681
K-1 Pre: Pretest Total		5.390	9.365	0.575	0.579
K-1 Pre: District & Pretest		5.989	10.536	0.568	0.590
1-3 All: Basic Two-Level		-7.269	4.36*	-1.668	0.096
1-3 All: School Curvature		-7.135	4.1**		
1-3 Pre: Basic Two-Level		-3.968	6.6**		
1-3 Pre: Pretest Total		-0.610	6.3**		
1-3 Pre: Propensity for Pre		-4.148	6.3**		
1-3 Pre: Pre Predicts Both		-1.097	6.3**		
1-3 Three-Level: No District		-6.549	7.29*	-0.899	0.381
1-3 Three-Level: District Main		-5.678	6.64*	-0.856	0.407
1-3 Three-Level: District EE Effects Only					
	District A	-23.714	13.6**		
	District B	0.439	13.9**		
	District C	5.014	12.68*	0.395	0.699
	District F	-6.740	18.7**		
	District H/I	-1.389	13.7**		
1-3 Three-Level: Program-District Interactions					
	District A	-23.713	18.0**		
	District B	17.642	21.0**		
	District C	-16.223	14.5**		
	District F	—	—		
	District H/I	0.686	16.5**		

* Estimated from t-statistic

** No t-statistic was calculated, so the standard error of the coefficient for the dummy variable as a predictor of initial status (intercept) is given.

For the reading subtest, program-district interactions were found to be statistically significant even after controlling for district main effects. An alternative model, allowing district differences among the EE schools but not among the IS schools, was found to fit the data nearly as well with fewer parameters. For this alternative model ("1-3 Three-Level: District EE Effects Only"), the difference between the EE program in each district and the overall estimate of the IS program at spring of first grade is generally small. The EE program in district A is estimated to have reading achievement scores nearly 24 points above the average for the IS schools, a difference that is not quite statistically significant, since it is less than twice the rough standard error of 13.6 points.

The full model of district differences for EE schools and separate district differences for IS schools ("1-3 Three-Level: Program-District Interactions") produces separate estimates for each program-district combination. The program differences are estimated separately for each district (district F has no one-program IS schools, so no program difference can be estimated for that district). The standard errors for these comparisons are relatively large because only the schools in each district can be used to compare the programs at spring of first grade. None of the differences approach statistical significance. Note that the estimated program differences in this table differ from those given in Table 198 because in this table the effect of the program difference in curvature is incorporated into the estimate.

Summary

The K-1 and 1-3 estimates of the difference between the two programs for the reading subtest at spring of first grade are not statistically significant. The overall estimates of program differences indicate a small advantage for the early-exit program, but the difference is less than one standard error and therefore far from statistically significant. Some insight into the origin of the early-exit advantage can be obtained from the 1-3 analysis incorporating district differences for the EE program and the analysis incorporating separate district differences for the EE program and the IS program (the full district-by-program interaction model). In both of these analyses, the EE program in district A shows nearly a 24 point advantage over the IS programs as a group or over the IS program in district A at spring of first grade. This difference is not statistically significant but is likely to be the cause of the estimated early-exit advantage. The early-exit programs in the other districts have first grade reading achievement scores that are about the same, on average, as the immersion strategy programs. It should be emphasized that none of the program differences are significant at the .05 level.

Conclusions

What are the overall conclusions from the one-program schools analyses?

Perhaps the most important result of the analyses of one-program schools is that school-to-school differences are large even within the same district and program. Any program differences are small compared to the school differences. No doubt some of the differences among schools can be attributed to variations in the program implementation at the schools, but such operational program analyses go beyond the scope of this report.

Because of the large school variability and the small number of schools available for the analysis, it is difficult to assess district and program differences. Nonetheless, some statistically significant program and district differences were detected.

The students in the early-exit and immersion strategy programs were found to differ in some material ways. A propensity score was estimated for each student, as described in Chapter II, and included as a predictor in the analyses. The propensity score adjusts the estimated program effect to account for background variable differences between the students in the two programs. All of the program and district differences summarized in this section are estimated after adjusting for the propensity score (see Table 208 for a summary of results).

Table 208

Summary of Program Differences in
One-Program School HLM Analyses

		Initial Status	Growth Rate	Curvature
Mathematics (see Table 174)	ESS Differences+	3.179	-3.183	-3.657
	p-Value	.319	.146	.164
English Language (see Table 175)	ESS Differences	3.817	4.710	-14.086
	p-Value	.448	.223	.001**
English Reading (see Table 176)	ESS Differences	-5.896	3.251	-4.118
	p-Value	.163	.322	.283

\+ Expanded Scale Score Differences between IS and EE

The mathematics subtest shows essentially no differences in growth curves between the programs. The initial status, growth rate, and curvature are all virtually identical in the two programs. There are marginally significant differences among the districts. At spring of first grade, district F has relatively high mathematics achievement and district B relatively low mathematics achievement, with the other three districts in between. Districts A and H/I show the highest growth rates, with district F next highest and districts B and C lowest.

For the language subtest, the immersion strategy programs have a higher status by 8 to 12 points at spring of first grade in the K-1 analyses, but the difference is not statistically significant. In the 1-3 analyses, the program difference at spring of first grade is much smaller, varies in direction, and is not statistically significant. The language growth rates for the two programs are not significantly different. However, there is substantially greater negative curvature in the immersion strategy program than in the early-exit program. Thus the immersion strategy schools have higher language achievement in second grade but have lower growth from second grade to third grade. The two programs are not significantly different in spring of third grade.

ERIC

For the language subtest, there are significant district differences in spring of first grade and in growth rate. In spring of first grade, districts A and F have the highest language achievement test scores. District H/I is next highest, followed by district B and then district C. Although the overall test for district differences in initial status was statistically significant, closely ranked districts were not significantly different from each other. For the growth rate from first grade to third grade, districts F and H/I were highest, district A next highest, and districts B and C lowest.

For the reading subtest, early-exit students had higher scores in spring of first grade by up to 7 points, but the difference is not statistically significant. The programs do not differ in second grade or third grade, and there are no overall program differences in growth rates or curvature. Although the test for district main effects was not significant, the test for district-program interactions was marginally significant. This implies that the program differences vary by district.

The district-program interactions for the reading subtest led to the consideration of an alternative model in which the early-exit schools exhibit district differences but the immersion strategy schools do not. This alternative model is preferable to the full district-program interaction model because it fits the reading test scores nearly as well with fewer parameters. The source of the interactions and the district differences within the early-exit program is in part due to the early-exit program in district A. Except for the early-exit program in district A, all of the early-exit and immersion strategy programs appear to have comparable initial status and growth rate on the reading subtest. The early-exit program in district A has substantially higher reading achievement at spring of first grade. The other source of the interactions appears to be the early-exit programs in districts B and C. The reading growth rates in the early-exit schools in districts B and C are lower than in the other districts and lower than the growth rates in the immersion strategy schools.

ERIC

The alternative model of district differences among the early-exit schools but not among the immersion strategy schools was tested on the mathematics and language subtests. It was found to fit the data about as well as the model of district main effects. The district main effects model says that districts have overall differences in achievement levels that apply equally to both programs. This model did not fit the mathematics and language data any better or any worse than the alternative model, which says the early-exit programs have differences in achievement across the districts but the immersion strategy programs do not have differences across the districts. As noted above, the alternative model did fit the reading data better than the district main effects model.

The results of these one-program schools analyses were found to be insensitive to the analytic method and to the covariates included as predictors of achievement. Estimated program differences at spring of first grade were similar in the K-1 analyses, the basic 1-3 analyses (with school treated as a fixed effect), and the 1-3 analyses using the three-level hierarchical linear model (with school treated as a random effect). The estimated program differences in growth rate and curvature in the 1-3 analyses were similar whether school was treated as a fixed effect or as a random effect. The estimated program effects were even found not to be affected by the inclusion of the propensity score, which implies that the students in the two programs are not very disparate.

The one sensitivity in these analyses of one-program schools involved the reduction to students with pretest scores available. The elimination of students without pretest scores available introduced a program difference in growth rate for mathematics and language and eliminated the program difference in curvature for language. (It had essentially no effect on the estimated program differences for reading.) The differences in the analyses are probably attributable to the sizable school differences. The elimination of several schools and school groups from the pretest analyses may have destroyed the relative similarity of the immersion strategy schools and the early-exit schools on the measure of school socioeconomic status (see Chapter II). Moreover, the introduced program

ERIC

differences in growth rates are not larger than the estimated differences for many of the individual schools and school groups. A more complete evaluation of the program difference in the mathematics and language growth rates among these students with pretest scores available would require additional analyses.

Summary

Almost no program differences were found among the one-program schools. The immersion strategy schools exhibited significantly greater curvature on the language subtest. Thus the immersion strategy schools had higher language achievement than the early-exit schools at second grade but there was no significant difference at third grade. This finding of a temporary boost in achievement for the students in immersion strategy schools is similar to the results for the language and reading subtests for the two-program schools. For the two-program schools, however, the temporary boost appears at first grade for the language subtest and at second grade for the reading subtest.

Some district differences were found among the one-program schools. Where there were differences, districts B and C tended to have lower status at first grade and lower growth rates; districts A, F, and H/I generally had higher status at first grade and higher growth rates. One of these three "high" districts was usually not as high as the other two, but not always the same one. Comparing this pattern of district differences with the patterns of instructional English described in Chapter I is suggestive of a relationship. The early-exit program in district F had the highest levels of English use, with districts A and H/I next highest. Districts B and C had the lowest use of English for instruction. It is possible that these variations in program implementation explain the significant district differences for the language subtest and, especially, the district-program interactions for the reading subtest.

V. THE INTRA-PROGRAM ANALYSES OF LATE-EXIT TRANSITIONAL BILINGUAL EDUCATION PROGRAMS

Introduction

What was the purpose of the late-exit analyses?

The analyses of students in the late-exit program were designed to compare the patterns of student growth across the districts, schools, and cohorts within the late-exit program. The one school in late-exit district D, the seven schools in late-exit district E, and the six schools in late-exit district G were compared within and between districts. Of special interest was the comparison of the kindergarten through third grade (K-3) cohort of students and the third grade through sixth grade (3-6) cohort of students. All comparisons were made in all three areas of academic achievement, as measured in English: mathematics, language arts, and reading.

To date, little is known about the academic growth of limited-English-proficient students over time. The information provided here will be useful to educators and researchers with hypotheses about the growth of these students from first grade through sixth grade.

Why were the late-exit students analyzed separately?

The late-exit students were analyzed separately in large part because the districts with late-exit programs had implemented no alternative programs. This makes it impossible to compare students in the late-exit program with students in other programs while controlling for district differences. The other main reason for separating the analyses of the late-exit students is that the immersion strategy and early-exit students were not studied beyond fourth grade, but a cohort of late-exit students was studied from third grade through sixth grade.

What statistical methods were used in the late-exit analyses?

The late-exit analyses were divided into two major parts, just as the IS/EE analyses were: a set of K-1 analyses using Analysis of Covariance (ANCOVA) and a set of analyses using individual growth curves. In the K-1 analyses, the schools were treated as random effects, so the HLM computer program was used to perform the ANCOVA. In the growth curve analyses, a two-level hierarchical linear model was estimated using the HLM program: each student had an individual growth curve estimated and the intercept and slope of these growth curves were then predicted using student attributes. These growth curve analyses include growth curves for first grade through third grade estimated using the K-3 cohort of students and growth curves for third grade through sixth grade using the 3-6 cohort of students. These analyses are therefore referred to as 1-6 analyses.

How similar are the three late-exit districts?

The three late-exit districts were found to have similar socioeconomic status at the school level (see Chapter II). However, the three late-exit districts serve very different populations. Late-exit district D is the only study district located in Florida, where the Spanish-speaking population includes many Cubans. Late-exit district E, in New York, serves a Spanish-speaking population that is largely Puerto Rican. Late-exit district G is located in southern California, where the Spanish-speaking population includes many students of Mexican descent.

As described in Chapter I, large differences were found among the three late-exit districts in the pattern of instructional English. The pattern of the amount of English used for instruction in district G was found to be closer to the model for an early-exit program than the model for a late-exit program. District D and district E showed a pattern of instructional English more faithful to the late-exit name. Thus the districts differ considerably in the educational program.

ERIC

What constraints were placed on the late-exit analyses by the study design?

Only one school in district D is included in the study. This makes it impossible to assess the school-to-school variation within district D. With 42 students, however, the estimated school effects should be stable enough to permit treating school 30, the one district D school, separately in the late-exit analyses.

Late-exit district E had seven schools participating in the study, but only six had K-3 cohorts (see Chapter II). Although all study students had been in the late-exit program from the beginning of kindergarten, the eight study students in school 42 all entered the study in third grade. In addition, school 46 had only four students with first grade test scores.

Late-exit district G had six study schools, but school 61 had only five students and school 64 only three students. More importantly, the "spring first grade" test for district G was actually administered in fall of second grade. Therefore, district G students were expected to have higher "spring first grade" test scores than they would have obtained had they actually been tested in spring of first grade. Furthermore, fall kindergarten tests were not available for any district G students because that district entered the study in spring of the first year, making it impossible to compare district G with the other two late-exit districts at fall of kindergarten. For these two reasons, if district G students had higher scores in "spring first grade" than students in the other two late-exit districts, the difference in scores could not be attributed to differences in the program implementation. Any difference in "spring first grade" scores could be due to differences that existed in fall of kindergarten or to the six-month difference in test date. Accordingly, most analyses were performed with district G treated separately, either by excluding the students entirely or by providing for differences between district G and the other two late-exit districts.

ERIC

In the absence of true first grade test scores for district G and in the absence of any way to assess the comparability of district G students with the other two late-exit districts in fall kindergarten, it is questionable to include district G in analyses that use "spring first grade" test scores. The analyses presented here include district G students only when doing so does not distort the results obtained when district G students are omitted.

To assess differences between the district E schools as a group and school 30, the single district D school, a measure of variability is needed. Because there is no information about the school-to-school variability in district D, it is necessary to rely on the school-to-school variability in district E. That is, school 30 is compared to the schools in district E to evaluate whether the achievement in school 30 is reasonably similar to the distribution of achievement among the district E schools.

Did the late-exit students exhibit curvature in growth?

In an extensive series of analyses not presented in detail here, separate growth curve models were estimated for first through third grade using the K-3 cohort and for third through sixth grade using the 3-6 cohort. These models were used to develop a set of predictors of the intercept and slope of the individual growth curves. They were also used to evaluate the curvature in growth from first through third grade and from third through sixth grade.

Substantial negative curvature was found for the immersion strategy and early-exit students in the 1-3 analyses reported in Chapter III and Chapter IV. That is, the growth was found to flatten, as expected from the national norms. For some of those IS/EE analyses, the amount of curvature varied by program.

In contrast, for the late-exit students in districts D and E the amount of curvature in growth within each cohort (cohorts K-3 and 3-6) was generally small. (District G was excluded from these analyses because of the problem with the "spring first grade" test score.) Where the curvature within cohort was significantly different from zero it was found to be positive: the students exhibited increasing growth rates, not decreasing growth rates. However, the 1-3 growth rate (estimated using the K-3 cohort) was found to be much greater than the 3-6 growth rate (estimated using the 3-6 cohort). That is, between cohorts the students exhibited decreasing growth rates, in accordance with expectations based on national norms.

Table 209 summarizes the estimated curvature coefficients for the two cohorts and all three subtests. For the K-3 cohort, the estimated curvature from spring of first grade through spring of third grade is negligible for the mathematics and reading subtests and does not approach statistical significance. For the language subtest, the estimated 1-3 curvature is significantly positive. For the 3-6 cohort, the estimated 3-6 curvature is not large for any of the three subtests. The negative curvature for the language subtest is not statistically significant at the .05 level. The positive curvature is statistically significant for mathematics and reading but does not exceed about 5.3 points; that is, the growth rate accelerates, increasing 5.3 points per year. A curvature coefficient of 5.3 points implies a growth rate about 16 points higher in sixth grade than in third grade.

Table 209

Late-Exit 1-3 and 3-6 Curvature Estimates

Description		Parameter	Std Err	t-Stat	p-Value
Mathematics	Curvature 1-3	1.942	2.736	0.710	0.479
Language	Curvature 1-3	11.166	4.043	2.762	0.007
Reading	Curvature 1-3	2.179	3.234	0.674	0.502
Mathematics	Curvature 3-6	4.371	1.783	2.452	0.016
Language	Curvature 3-6	-2.572	1.450	-1.774	0.079
Reading	Curvature 3-6	5.308	1.702	3.118	0.002

The growth rate analyses presented in this chapter do not include within-cohort curvature. They do include between-cohort curvature in the sense that the 1-3 growth rates and the 3-6 growth rates are allowed to differ. That is, the individual student growth curves are actually straight lines, and the slopes of the lines for the K-3 cohort of students are modeled to be steeper than the slopes of the lines for the 3-6 cohort of students.

The decision to present the models without curvature within cohort was based on several factors. If the true growth within cohort is quadratic, fitting a straight line to the test scores gives an accurate estimate of the average growth. The curvature within cohort, if it exists, is small and positive (see Table 209). The curvature between cohorts is substantially larger and negative (see Table 250, Table 251, and Table 252 and the accompanying discussion later in this chapter). Thus it seems likely that if students had been studied longitudinally from kindergarten through sixth grade, the overall pattern of curvature in growth would have been negative.

In addition, combining the positive curvature within the 1-3 and 3-6 cohorts with the negative curvature across the 1-3 and 3-6 cohorts would have produced a complex model. That doubly-curved model of student growth would have been difficult to interpret and would be unstable with the relatively small number of students available for analysis. With no students studied longitudinally across the third grade boundary, it is inappropriate to reach conclusions about shape of student growth at third grade: the implied "cusp" of the doubly-curved model at third grade is quite possibly artifactual.

The effect of fitting straight-line individual growth curves is to relegate any within-student curvature to the estimated within-student variability. Table 210 shows the estimated within-student standard error for the models with and without curvature for both cohorts and all three subtests. As expected from the estimated curvature coefficients in Table 209, the greatest effect is for the two coefficients that were significant at the .01 level: the 1-3 curvature for the language subtest and the 3-6 curvature for the reading subtest. In both cases the increase in standard error is less than two percent, clearly a negligible change. Thus, while there is some curvature in growth, these analyses indicate that the use of straight-line rather than quadratic growth curves will have a negligible effect on the results.

Table 210

Late-Exit 1-3 and 3-6 Curvature Effect on Within-Student Variability

Description		Estimated Within-Student Standard Error With Curvature	Without Curvature
Mathematics	1-3	21.94	21.80
Language	1-3	32.52	33.11
Reading	1-3	26.05	26.00
Mathematics	3-6	29.42	29.64
Language	3-6	24.14	24.28
Reading	3-6	28.61	29.12

Results of the K-1 Analyses for Mathematics

What were the results of the basic K-1 analyses for the mathematics subtest?

The purpose of the K-1 analyses for mathematics was to establish that students in the different late-exit districts had comparable skills at the end of kindergarten. The first K-1 analysis used the HLM program to fit an overall average to assess the amount of variation among the 13 schools with K-1 cohorts on the mathematics subtest (see Table 211). Highly significant school variation was found, as expected. However, the within-school variation was found to be fairly uniform among the 13 schools (chi-square = 15.02 with 12 d.f., p = .240). This means that the assumption that the variability of the students is the same within each of the 13 schools cannot be rejected.

Table 211

Late-Exit K-1 Analysis for Mathematics
All Students; No Predictors

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	248.256	5.193	47.809	0.000

The second step was to include a variable with the value one for the six schools in district G and zero for the other seven schools (see Table 212). This analysis indicates the expected higher scores for district G schools. The parameter for DISTG represents the difference between district G schools and the other schools. The parameter estimate indicates district G students score significantly higher (difference = 22.966, p = .023). As noted above, this estimated difference is undoubtedly affected by the fact that the "spring first grade" tests in district G were actually administered in fall of second grade. Although including the district G dummy variable reduces the estimated school variability from a standard error of 16.6 to a standard error of 12.9, the remaining school variability is still highly significant.

Table 212

Late-Exit K-1 Analysis for Mathematics
All Students; District G Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	238.717	5.557	42.959	0.000
Spring 1st	DISTG	22.966	8.691	2.643	0.023

The third step was to add a dummy variable for indicating school 30, the one district D school (see Table 213). The parameter estimate for BASE now represents the average mathematics score among the six district E schools. The parameter estimate for DISTG indicates the difference between the district G schools as a group and the district E schools, and

ERIC

the parameter estimate for SCHOOL30 indicates the difference between school 30 and the district E schools. Both the DISTG and SCHOOL30 parameters are statistically significant (difference = 28.041, p = .003 for DISTG; difference = 29.734, p = .021 for SCHOOL30). That is, both the district G schools as a group and the single district D school have significantly higher mathematics scores than the district E schools as a group. Furthermore, by estimating a separate average for school 30 as well as district G, the estimate of school variability drops considerably, to a standard error of 8.8. Although an approximate test for school differences is still significant at the .05 level, it is not significant at the .01 level (chi-square = 22.551 with 10 d.f., p = .013). This means that most of the school-to-school variability on the mathematics subtest can be attributed to differences among the three late-exit districts.

Table 213

Late-Exit K-1 Analysis for Mathematics
All Students; School 30 and District G Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	233.933	4.641	50.410	0.000
Spring 1st	DISTG	28.041	7.097	3.951	0.003
Spring 1st	SCHOOL30	29.734	10.910	2.725	0.021

Summary

The schools in district G and the one school in district D have significantly higher mathematics subtest scores at "spring first grade" than the schools in district E. With no covariates included, most of the school-to-school variability in mathematics is explained by district differences. Because the "spring first grade" tests in district G were actually administered in fall of second grade, district G students were expected to have relatively high scores.

How were the K-1 analyses for mathematics affected by including covariates?

For the late-exit districts, seven background variables were found to be predictive of achievement on the three subtests. Of these seven background variables, five are the ones employed in the IS/EE analyses reported in Chapters III and IV: EDAVG (parents' educational average), PRESCHY (the student attended preschool), FEMALE (the student is a girl), BOOKSHM (the number of books in the home), and ANYEPTOP (the parents use any English when speaking to each other). The two additional background variables are OCAVG (average occupational status of the parents) and AGEMOTH (mother's age, in three categories). For additional information on all of the background variables, see Chapter II.

Of these seven background variables plus the number of absences in first grade (ABS1), the key covariates for predicting the "spring first grade" mathematics subtest score were found to be ABS1 and BOOKSHM. A set of three analyses was performed to evaluate the effect of these covariates on the district differences uncovered in the basic K-1 analyses for the mathematics subtest. Table 214 shows the model with the key covariates included but no district differences accommodated. Table 215 shows the effect of adding the DISTG variable, and Table 216 shows the effect of including both the DISTG and SCHOOL30 variables.

Table 214

Late-Exit K-1 Analysis for Mathematics
All Students; Key Covariates

Predicted	Predictor	Parameter	Std.Err	t-Stat	p-Value
Spring 1st	BASE	247.558	6.025	41.088	0.000
Spring 1st	ABS1	-0.571	0.194	-2.944	0.004
Spring 1st	BOOKSHM	3.774	1.505	2.508	0.013

Table 215

Late-Exit K-1 Analysis for Mathematics
All Students; District G and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	237.747	6.677	35.606	0.000
Spring 1st	DISTG	21.397	8.226	2.601	0.025
Spring 1st	ABS1	−0.522	0.194	−2.689	0.008
Spring 1st	BOOKSHM	4.011	1.497	2.680	0.008

Table 216

Late-Exit K-1 Analysis for Mathematics
All Students; School 30, District G, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	234.529	6.635	35.346	0.000
Spring 1st	DISTG	25.072	7.871	3.185	0.010
Spring 1st	SCHOOL30	20.801	12.511	1.663	0.127
Spring 1st	ABS1	−0.481	0.195	−2.463	0.015
Spring 1st	BOOKSHM	3.624	1.508	2.404	0.017

A comparison of these tables with the corresponding tables without covariates (Table 211 through Table 213) shows that the district G schools are still significantly different from the others, whether or not SCHOOL30 is included (p = .025 and p = .010). After adjusting for these two covariates, however, the difference between school 30 and the schools in district E is reduced from 29.734 to 20.801 and is no longer statistically significant (p = .127). This implies that part of the difference between school 30 and the district E schools is explained by these two covariates. Because the number of absences is unlikely to vary considerably by district, it is probably the inclusion of BOOKSHM that produces this result.

The inclusion of the two key covariates has the expected effect of reducing the within-school variability, from a standard error of about 29

ERIC

to a standard error of about 28. Including the two key covariates also has the effect of reducing the between-school variability in the model without district variables (see Table 214) and in the model with DISTG only (see Table 215). However, the estimated between-school variability increases slightly in the model with DISTG and SCHOOL30 (from a standard error of 8.8 to a standard error of 10.3). This is partly a consequence of the decreased within-school variability: more of the overall variability can now be attributed to between-school differences. The approximate test for school-level variability is now significant (chi-square = 27.919 with 10 d.f., p = .002).

Table 217 through Table 219 show the corresponding analyses with all eight covariates included. Adding the other six variables has only a negligible effect on the within-school variability. The district G schools are still significantly different (p = .043 and p = .031), and school 30 is again not significantly different from the district E schools (p = .306). In other words, adding the other six covariates does not appreciably affect the results from the models with two key covariates.

Table 217

Late-Exit K-1 Analysis for Mathematics
All Students; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	229.265	13.486	17.001	0.000
Spring 1st	ABS1	-0.504	0.200	-2.523	0.012
Spring 1st	EDAVG	0.381	0.750	0.507	0.613
Spring 1st	PRESCHY	8.623	4.690	1.839	0.067
Spring 1st	FEMALE	-2.269	3.941	-0.576	0.565
Spring 1st	OCAVG	1.759	1.226	1.435	0.153
Spring 1st	AGEMOTH	3.475	4.394	0.791	0.430
Spring 1st	BOOKSHM	3.156	1.602	1.969	0.050
Spring 1st	ANYEPTOP	-1.897	5.541	-0.342	0.733

Table 218

Late-Exit K-1 Analysis for Mathematics
All Students; District G and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	218.638	14.132	15.471	0.000
Spring 1st	DISTG	19.941	8.732	2.284	0.043
Spring 1st	ABS1	-0.453	0.200	-2.264	0.025
Spring 1st	EDAVG	0.621	0.759	0.818	0.414
Spring 1st	PRESCHY	6.495	4.821	1.347	0.179
Spring 1st	FEMALE	-2.224	3.933	-0.565	0.573
Spring 1st	OCAVG	1.738	1.220	1.425	0.156
Spring 1st	AGEMOTH	3.792	4.374	0.867	0.387
Spring 1st	BOOKSHM	3.249	1.598	2.033	0.043
Spring 1st	ANYEPTOP	-1.852	5.527	-0.335	0.738

Table 219

Late-Exit K-1 Analysis for Mathematics
All Students; School 30, District G, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	217.857	14.120	15.429	0.000
Spring 1st	DISTG	22.599	8.979	2.517	0.031
Spring 1st	SCHOOL30	15.091	13.967	1.080	0.306
Spring 1st	ABS1	-0.430	0.201	-2.134	0.034
Spring 1st	EDAVG	0.537	0.762	0.705	0.482
Spring 1st	PRESCHY	5.992	4.847	1.236	0.218
Spring 1st	FEMALE	-2.182	3.935	-0.555	0.579
Spring 1st	OCAVG	1.647	1.222	1.347	0.179
Spring 1st	AGEMOTH	3.501	4.382	0.799	0.425
Spring 1st	BOOKSHM	3.105	1.603	1.936	0.054
Spring 1st	ANYEPTOP	-1.964	5.530	-0.355	0.723

Summary

The purpose of the K-1 analyses for mathematics was to examine the comparability of students across the different late-exit districts at spring of first grade. Adding two key covariates (number of first grade absences and the number of books in the home) does not change the finding

ERIC

that the schools in district G have significantly higher mathematics subtest scores at "spring first grade" than the schools in district E. The two covariates have the effect of reducing the estimated difference between school 30 in district D and the district E schools from about 30 points to about 21 points and the difference is no longer statistically significant. Although much of the school-to-school variability in mathematics is explained by district differences after including the two key covariates, the remaining school variability is statistically significant. Adding additional covariates has no marked effects.

How were the K-1 analyses for mathematics affected by controlling for pretest?

Because district G entered the study in spring, none of the district G students have fall kindergarten test scores available. That is, district G cannot be included in the pretest analyses. However, nearly all students in all six district E schools can be included along with half of the students from school 30, the one school in district D. Table 220 shows the basic key-covariates model for the mathematics subtest for the students with pretest scores available. Table 221 shows the same model with SCHOOL30 included. The estimated difference in mathematics achievement between school 30 and the schools in district E is 20.863, remarkably close to the corresponding value in Table 216, 20.801. Neither value is statistically significant.

Table 220

Late-Exit K-1 Analysis for Mathematics
Pretest Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	240.185	7.665	31.335	0.000
Spring 1st	ABS1	-0.563	0.223	-2.524	0.012
Spring 1st	BOOKSHM	2.954	1.734	1.704	0.090

Table 221

Late-Exit K-1 Analysis for Mathematics
Pretest Students; School 30 and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	237.111	7.731	30.669	0.000
Spring 1st	SCHOOL30	20.863	15.793	1.321	0.278
Spring 1st	ABS1	-0.535	0.224	-2.393	0.018
Spring 1st	BOOKSHM	2.724	1.739	1.566	0.119

The reduction to the students who have pretest scores available and the consequent elimination of district G has had the effect of reducing the estimated coefficient of BOOKSHM and making the variable nonsignificant (p = .090 and p = .119). The within-school variability is also reduced, from a standard error of about 28 to a standard error of 25.5 (finding not presented in a table). The between-school variability is little changed and remains highly statistically significant.

Table 222 and Table 223 show the effect of adding the mathematics pretest to the basic model and to the model with SCHOOL30, respectively. For these two districts, the mathematics pretest was the best predictor of spring first grade mathematics achievement, so the results from using the other pretest scores are not presented. The estimated difference between school 30 and the district E schools is reduced to about 15 and remains nonsignificant. The coefficients for ABS1 and BOOKSHM are both reduced and now neither is statistically significant. The within-school variability is further reduced, from a standard error of 25.5 to a standard error of about 24. Although the between-school variability is slightly reduced, to a standard error of about 12, it is highly statistically significant.

Table 222

Late-Exit K-1 Analysis for Mathematics
Pretest Students; Key Covariates and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	148.588	23.455	6.335	0.001
Spring 1st	ABS1	-0.386	0.215	-1.792	0.075
Spring 1st	BOOKSHM	1.519	1.678	0.905	0.367
Spring 1st	ANALPREM	0.566	0.139	4.089	0.000

Table 223

Late-Exit K-1 Analysis for Mathematics
Pretest Students; School 30, Key Covariates, and Mathematics Pretest

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	148.685	23.436	6.344	0.001
Spring 1st	SCHOOL30	14.965	14.079	1.063	0.336
Spring 1st	ABS1	-0.366	0.216	-1.690	0.093
Spring 1st	BOOKSHM	1.379	1.682	0.820	0.413
Spring 1st	ANALPREM	0.552	0.139	3.961	0.000

Summary

Reducing the students to those with pretest scores available (and therefore eliminating district G) does not change the main findings from the K-1 mathematics achievement analyses with the two key covariates included. The within-school variability is somewhat reduced for the pretest students in districts D and E compared to the value obtained using all students in districts D, E, and G. Students in school 30, the only school in district D, have higher mathematics achievement but the difference is not statistically significant. Adding the mathematics pretest as a predictor further reduces the within-school variability, as expected. It also has the effect of reducing the estimated difference between school 30 and the schools in district E from about 21 points to about 15 points, with the difference remaining not statistically significant.

What were the conclusions from the K-1 analyses for the mathematics subtest?

The objective of the K-1 analyses was to determine the comparability of students in the different late-exit groups on their mathematics skills at the end of first grade. For the mathematics subtest, the schools in district G have significantly higher mathematics subtest scores at "spring first grade" than the schools in district E. This is unsurprising because the "spring first grade" tests in district G were actually administered in fall of second grade. The estimate of the difference in mathematics achievement between district G and district E schools is not greatly affected by the inclusion of covariates. In the absence of any pretest scores for district G students, it is not possible to evaluate the extent to which district E and district G students were comparable on standardized tests upon entering kindergarten.

School 30, the only study school in district D, has significantly higher mathematics test scores in spring of first grade than the schools in district E when no covariates are included in the model. Including the number of absences in first grade (ABS1) and the number of books in the home (BOOKSHM) reduces the estimated difference between school 30 and the district E schools from about 30 ESS points to about 21 points and it becomes nonsignificant. Adding additional covariates, including the mathematics pretest, tends to reduce the difference still further and it remains nonsignificant.

With no covariates included, there is little school-to-school variability on mathematics not explained by district differences. With the addition of covariates, more of the overall variability can be attributed to between-school variability. The between-school differences remaining after controlling for district are highly significant. In other words, there are significant differences between the schools in the same district after controlling for the two key covariates. This finding parallels the finding of significant school differences within district and program among the one-program schools in the IS/EE districts (see Chapter IV).

Results of the K-1 Analyses for Language

What were the results of the basic K-1 analyses for the language subtest?

The K-1 analyses for English language skills were performed to determine if students across the three late-exit districts had comparable skills. The K-1 analyses for the language subtest parallel the K-1 analyses for the mathematics subtest. The model that included only an overall constant without allowing for district differences (see Table 224) showed highly significant between-school variation among the 13 schools. Unlike the mathematics subtest, the language subtest also showed significant differences in the within-school variation across the schools (chi-square = 41.38 with 12 d.f., p less than .0005). That is, some schools show more variability among their students than other schools. The differences in variability may be attributable to differences among classrooms (teacher background or style, or the classroom composition, for example) in the same school. Analyses to evaluate this possibility go beyond nominal program analyses and are therefore beyond the scope of this report.

Table 224

Late-Exit K-1 Analysis for Language
All Students; No Predictors

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	250.759	7.422	33.784	0.000

The inclusion of a dummy variable for district G (see Table 225) shows that the district G students have only slightly higher scores than

ERIC

the average for the other seven schools (difference = 8.401, p = .602). Estimating a separate level of language achievement at "spring first grade" for school 30, the only school in district D, reveals that school 30 exhibits much higher language achievement than average for either district E or district G (see Table 226). The estimated difference between school 30 and the average for district E schools is 61.169 (p = .004). Separating school 30 from the district E schools reveals that the average for the district E schools is about 16 points lower than for district G, but the difference is still not statistically significant (p = .156). That is, even though the "spring first grade" test in district G was in fact administered in fall of second grade, the average language achievement in district G schools is not significantly higher than in district E schools. The estimated difference between school 30 and the district G schools is 44.726 (calculated as 61.169 - 16.443). This is over 2.5 times as big as the standard error of the SCHOOL30 parameter and therefore may be considered large.

Table 225

Late-Exit K-1 Analysis for Language
All Students; District G Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	247.211	10.209	24.215	0.000
Spring 1st	DISTG	8.401	15.634	0.537	0.602

Table 226

Late-Exit K-1 Analysis for Language
All Students; School 30 and District G Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	237.664	7.052	33.699	0.000
Spring 1st	DISTG	16.443	10.720	1.534	0.156
Spring 1st	SCHOOL30	61.169	16.731	3.656	0.004

Estimating a separate average for school 30 as well as district G reduces the estimated school variability from a standard error of about 24 or 25 to a standard error of about 13.7. This between-school variability can be compared with the within-school variability of about 41.6. Although the between-school variability on the language subtest is still statistically significant (approximate chi-square = 24.783 with 10 d.f., $p = .006$), most of it has been explained by the high "spring first grade" achievement level in school 30.

Summary

The schools in district G have higher language subtest scores at "spring first grade" but the difference is not statistically significant. This is in spite of the fact that in district G the "spring first grade" tests were in fact administered in fall of second grade. The one school in district D has significantly higher language subtest scores at "spring first grade" than the schools in district E or district G.

How were the K-1 analyses for language affected by including covariates?

For the language subtest, the key covariates for predicting the "spring first grade" achievement score were found to be OCAVG (average occupational status of the parents) and BOOKSHM (the number of books in the home). Table 227 shows the analysis with these key covariates but without the DISTG or SCHOOL30 variables, Table 228 shows the analysis after adding DISTG, and Table 229 shows the analysis with both DISTG and SCHOOL30 included.

Table 227

Late-Exit K-1 Analysis for Language
All Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	216.069	9.735	22.196	0.000
Spring 1st	OCAVG	4.479	1.658	2.701	0.007
Spring 1st	BOOKSHM	8.803	2.133	4.127	0.000

Table 228

Late-Exit K-1 Analysis for Language
All Students; District G and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	209.692	11.490	18.250	0.000
Spring 1st	DISTG	13.993	13.332	1.050	0.316
Spring 1st	OCAVG	4.528	1.658	2.731	0.007
Spring 1st	BOOKSHM	8.977	2.139	4.198	0.000

Table 229

Late-Exit K-1 Analysis for Language
All Students; School 30, District G, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	205.095	10.247	20.014	0.000
Spring 1st	DISTG	19.999	11.087	1.804	0.101
Spring 1st	SCHOOL30	44.385	18.070	2.456	0.034
Spring 1st	OCAVG	4.123	1.659	2.485	0.014
Spring 1st	BOOKSHM	8.579	2.145	4.000	0.000

Compared to the models without covariates, the parameter estimates for DISTG are somewhat increased but still do not reach statistical significance. The parameter estimate for SCHOOL30 is reduced from 61.169 to 44.385 but is still statistically significant ($p = .034$). That is, the difference in spring first grade language achievement between school 30 and the schools in district E is reduced by adjusting for the two key covariates, OCAVG and BOOKSHM. The coefficients for the covariates scarcely change with the addition of the dummy variables, indicating that the within-district relationship between these covariates and achievement is similar to the between-district relationship. These facts imply that school 30 has higher-scoring students than the average school in district E, but the covariates explain much of the difference. Because of the six-month delay in testing district G students compared to the other two districts, no similar interpretation can be given to the higher language achievement scores (or lack thereof) in district G.

The addition of the key covariates reduces the within-school variability from a standard error of about 42 to a standard error of about 39.5. As was found for the mathematics subtest, including the two key covariates also has the effect of reducing the between-school variability in the model without district variables (see Table 227) and in the model with DISTG only (see Table 228) and increasing it slightly in the model with DISTG and SCHOOL30 (from a standard error of 13.7 to a standard error of 14.9). Again, this is partly a consequence of the decreased within-school variability.

The effect of adding the other six background variables is shown in Table 230 through Table 232. The district G schools are still not significantly different from the other schools in "spring first grade" language achievement and the estimates of within-school and between-school variability are little changed. However, the estimated difference between school 30 and the schools in district E is reduced from 44.385 to 39.341 and is now just short of statistical significance ($p = .054$). This confirms the trend that adding covariates reduces the estimated difference between school 30 and the district E schools.

Table 230

Late-Exit K-1 Analysis for Language
All Students; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	192.861	18.941	10.182	0.000
Spring 1st	ABS1	-0.088	0.283	-0.313	0.755
Spring 1st	EDAVG	1.710	1.061	1.612	0.108
Spring 1st	PRESCHY	5.579	6.629	0.842	0.401
Spring 1st	FEMALE	1.981	5.596	0.354	0.724
Spring 1st	OCAVG	3.566	1.738	2.052	0.041
Spring 1st	AGEMOTH	6.532	6.230	1.048	0.296
Spring 1st	BOOKSHM	7.648	2.273	3.365	0.001
Spring 1st	ANYEPTOP	1.131	7.864	0.144	0.886

Table 231

Late-Exit K-1 Analysis for Language
All Students; District G and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	184.024	20.410	9.016	0.000
Spring 1st	DISTG	15.550	13.317	1.168	0.267
Spring 1st	ABS1	-0.039	0.285	-0.137	0.891
Spring 1st	EDAVG	1.953	1.080	1.808	0.072
Spring 1st	PRESCHY	3.456	6.856	0.504	0.615
Spring 1st	FEMALE	2.053	5.588	0.367	0.714
Spring 1st	OCAVG	3.590	1.736	2.068	0.040
Spring 1st	AGEMOTH	6.715	6.224	1.079	0.282
Spring 1st	BOOKSHM	7.768	2.272	3.418	0.001
Spring 1st	ANYEPTOP	0.931	7.855	0.118	0.906

Table 252

Late-Exit K-1 Analysis for Language
All Students; School 30, District G, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	181.213	19.735	9.188	0.000
Spring 1st	DISTG	21.301	11.900	1.790	0.104
Spring 1st	SCHOOL30	39.341	18.005	2.185	0.054
Spring 1st	ABS1	0.006	0.285	0.022	0.982
Spring 1st	EDAVG	1.773	1.081	1.640	0.103
Spring 1st	PRESCHY	2.708	6.887	0.393	0.695
Spring 1st	FEMALE	2.017	5.592	0.361	0.718
Spring 1st	OCAVG	3.289	1.734	1.896	0.059
Spring 1st	AGEMOTH	6.438	6.219	1.035	0.302
Spring 1st	BOOKSHM	7.435	2.277	3.265	0.001
Spring 1st	ANYEFTOP	0.677	7.855	0.086	0.932

Summary

A series of sensitivity analyses was completed to ensure that the results of the main K-1 analyses were not influenced by factors other than program. Adding covariates does not change the finding that the schools in district G do not have significantly higher language subtest scores at "spring first grade" than the schools in district E despite being tested six months later. Adding covariates reduces the difference between school 30 in district D and the district E schools. The estimated difference is reduced from about 61 points (p = .004) to about 44 points (p = .034) by the addition of the two key covariates, occupational status of the parents and books in the home. The difference is further reduced to about 39 points (p = .054) by the addition of six other background variables. The addition of covariates somewhat reduces the within-school variability and increases the between-school variability. The between-school variability remains statistically significant.

How were the K-1 analyses for language affected by controlling for pretest?

As was done for the mathematics subtest, the effect of including the pretest was evaluated using the model including key covariates. District G students do not have fall kindergarten test scores and therefore cannot be included in the pretest analyses. For the language subtest, the best predictor of spring first grade achievement was found to be ANALTOBE, the sum of the mathematics and language subtest scores. Table 233 through Table 236 show the estimated models of language achievement for the pretest students.

Table 233

Late-Exit K-1 Analysis for Language
Pretest Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	220.035	12.637	17.412	0.000
Spring 1st	OCAVG	3.097	1.967	1.574	0.117
Spring 1st	BOOKSHM	6.957	2.318	3.001	0.003

Table 234

Late-Exit K-1 Analysis for Language
Pretest Students; School 30 and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	218.907	10.195	21.471	0.000
Spring 1st	SCHOOL30	48.658	14.319	3.398	0.043
Spring 1st	OCAVG	2.037	1.948	1.046	0.297
Spring 1st	BOOKSHM	6.141	2.312	2.656	0.009

Table 235

Late-Exit K-1 Analysis for Language
Pretest Students; Key Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	179.618	23.101	7.775	0.000
Spring 1st	OCAVG	2.771	1.953	1.419	0.157
Spring 1st	BOOKSHM	6.219	2.327	2.673	0.008
Spring 1st	ANALTOBE	0.126	0.061	2.049	0.042

Table 236

Late-Exit K-1 Analysis for Language
Pretest Students; School 30, Key Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	179.308	21.569	8.313	0.000
Spring 1st	SCHOOL30	44.582	11.472	3.886	0.012
Spring 1st	OCAVG	1.456	1.913	0.761	0.448
Spring 1st	BOOKSHM	5.181	2.300	2.252	0.025
Spring 1st	ANALTOBE	0.129	0.059	2.192	0.029

For these students with pretest scores, which necessarily excludes all district G students, the coefficient of OCAVG as a predictor of the language pretest score is reduced and is no longer statistically significant. However, the number of books in the home is still a statistically significant predictor of language achievement at spring of first grade. The within-school variability is reduced, from a standard error of about 39.5 to a standard error of about 34. In the models with SCHOOL30 included, the between-school variability is substantially reduced, from a standard error of 14.9 to a standard error of 10.5 (without the pretest as a predictor) or 6.7 (with the pretest as a predictor). The estimated difference between school 30 and the district E schools remains over 40 points and is marginally statistically significant ($p = .043$ and $p = .012$).

Summary

The main K-1 analyses were completed without including pretest data. This was done to maximize the number of students in the HLM analyses, thereby increasing the statistical power of the analyses. A sensitivity analysis was done by effecting an HLM analysis which included pretest data as a covariate. If a comparison of the sensitivity test and the original HLM results indicated similar findings, this would further strengthen the original results. Reducing the students to those with pretest scores available (and therefore eliminating district G) does not change the main findings from the K-1 language achievement analyses. The within-school variability is somewhat reduced and the between-school variability markedly reduced compared to the value obtained using all students in districts D, E, and G; the between-school variability is further reduced by including the pretest total as a predictor. School 30, the only school in district D, has significantly higher language achievement.

What were the conclusions from the K-1 analyses for the language subtest?

The K-1 analyses for English language were done to verify that students in all late-exit districts had comparable skills at the end of first grade. For the language subtest, the schools in district G have higher language subtest scores at "spring first grade" than the schools in district E but the difference is small and never statistically significant. School 30, the only study school in district D, has significantly higher language test scores in spring of first grade than the schools in district E. When no covariates are included in the model, the difference is over 60 points. With the parents' occupational status (OCAVG) and the number of books in the home (BOOKSHM) included in the model, the estimated difference between school 30 and the district E schools is reduced to about 44 points and is marginally statistically significant. Adding additional covariates, including the pretest total score, changes the estimated difference by no more than about five points in either direction, with the statistical significance remaining marginal ($p = .012$ to $p = .054$).

ERIC

There is considerable school-to-school variability on the language subtest (the estimated standard error is around 24 ESS points), but much of it is explained by district differences (reducing the standard error to 13.7 points). Among the students with pretest scores available, which necessarily excludes district G, including SCHOOL30 and the pretest total score as predictors reduces the between-school variability to a standard error of 6.7 ESS points.

Results of the K-1 Analyses for Reading

What were the results of the basic K-1 analyses for the reading subtest?

The K-1 analyses of reading achievement are similar to the analyses for the mathematics and language subtests. The analytic objective was to determine whether students in all of the late-exit districts had comparable reading skills at the end of first grade. The model that included only an overall constant without allowing for district differences (see Table 237) showed highly significant between-school variation among the 13 schools. Like the language subtest, the reading subtest showed significant differences in the within-school variation across the schools (chi-square = 43.84 with 12 d.f., p less than .0005). Again, the differences in variability within schools may be attributable to classroom differences but such operational analyses are beyond the scope of this report.

Table 237

Late-Exit K-1 Analyses for Reading
All Students, No Predictors

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	241.726	6.966	34.699	0.000

The inclusion of a dummy variable for district G (see Table 238) shows that the district G students have significantly higher scores than the average for the other seven schools (difference = 32.734, p = .014). Adding the dummy variable for school 30 (see Table 239) shows that school 30 and the district G schools have reading achievement at "spring first grade" about 40 points higher than the district E schools (p = .001 or less). The difference between school 30 and the district G schools is less than 6 points despite the fact that the "spring first grade" tests in district G were administered six months later.

Table 238

Late-Exit K-1 Analysis for Reading
All Students; District G Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	228.050	7.156	31.[illegible]70	0.000
Spring 1st	DISTG	32.734	11.144	2.937	0.014

Table 239

Late-Exit K-1 Analysis for Reading
All Students; School 30 and District G Only

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	222.234	4.327	51.354	0.000
Spring 1st	DISTG	37.899	6.932	5.467	0.000
Spring 1st	SCHOOL30	43.290	9.467	4.573	0.001

Estimating average reading achievement separately for the three late-exit districts reduces the estimated school variability from a standard error of about 22.6 to a standard error of about 6.4. This between-school variability is small compared with the within-school variability of about 35.6. After accounting for district differences, an approximate test of this between-school variability on the reading subtest is not statistically significant (chi-square = 16.262 with 10 d.f., p = .092).

Summary

The objective of these analyses was to determine whether students across the late-exit districts had comparable reading skills at the end of first grade. The only school in district D and the schools in district G have significantly higher reading subtest scores at "spring first grade" than the schools in district E. Essentially all of the between-school variation in "spring first grade" reading achievement is explained by district differences.

How were the K-1 analyses for reading affected by including covariates?

For the reading subtest, the same key covariates were used for predicting achievement as were used for the language subtest: OCAVG (average occupational status of the parents) and BOOKSHM (the number of books in the home). Table 240 through Table 242 show the analyses with these two covariates included. The results are similar to the results from the models without covariates, except that the estimated difference between school 30 and the district E schools is considerably reduced (from 43.290 to 27.620) although it is still significant ($p = .013$). Adding the other six background variables as covariates (see Table 243 through Table 245) has only minor effects. The estimated difference in reading achievement between school 30 and the schools in district E is further reduced, to 21.757, but remains statistically significant ($p = .040$).

Table 240

Late-Exit K-1 Analysis for Reading
All Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	211.649	9.118	23.213	0.000
Spring 1st	OCAVG	3.962	1.419	2.792	0.006
Spring 1st	BOOKSHM	7.564	1.826	4.143	0.000

ERIC

Table 241

Late-Exit K-1 Analysis for Reading
All Students; District G and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	193.600	8.198	23.617	0.000
Spring 1st	DISTG	37.418	8.358	4.477	0.001
Spring 1st	OCAVG	4.303	1.391	3.093	0.002
Spring 1st	BOOKSHM	8.290	1.797	4.613	0.000

Table 242

Late-Exit K-1 Analysis for Reading
All Students; School 30, District G, and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	192.599	7.094	27.150	0.000
Spring 1st	DISTG	40.714	6.456	6.306	0.000
Spring 1st	SCHOOL30	27.620	9.159	3.016	0.013
Spring 1st	OCAVG	3.872	1.387	2.792	0.006
Spring 1st	BOOKSHM	7.728	1.800	4.294	0.000

Table 243

Late-Exit K-1 Analysis for Reading
All Students; All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	192.916	16.276	11.853	0.000
Spring 1st	ABS1	-0.243	0.241	-1.008	0.315
Spring 1st	EDAVG	0.276	0.905	0.305	0.761
Spring 1st	PRESCHY	11.793	5.659	2.084	0.038
Spring 1st	FEMALE	2.710	4.754	0.570	0.569
Spring 1st	OCAVG	3.422	1.479	2.314	0.022
Spring 1st	AGEMOTH	8.163	5.300	1.540	0.125
Spring 1st	BOOKSHM	6.751	1.933	3.493	0.001
Spring 1st	ANYEPTOP	4.347	6.684	0.650	0.516

ERIC

Table 244

Late-Exit K-1 Analysis for Reading
All Students; District G and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	172.437	16.125	10.693	0.000
Spring 1st	DISTG	33.073	8.342	3.964	0.002
Spring 1st	ABS1	-0.169	0.236	-0.715	0.475
Spring 1st	EDAVG	0.854	0.901	0.948	0.344
Spring 1st	PRESCHY	8.799	5.756	1.529	0.128
Spring 1st	FEMALE	2.523	4.722	0.534	0.594
Spring 1st	OCAVG	3.598	1.453	2.477	0.014
Spring 1st	AGEMOTH	9.063	5.209	1.740	0.083
Spring 1st	BOOKSHM	7.207	1.909	3.776	0.000
Spring 1st	ANYEPTOP	4.266	6.620	0.644	0.520

Table 245

Late-Exit K-1 Analysis for Reading
All Students; School 30, District G, and All Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	174.146	15.536	11.209	0.000
Spring 1st	DISTG	36.476	7.354	4.960	0.001
Spring 1st	SCHOOL30	21.757	9.208	2.363	0.040
Spring 1st	ABS1	-0.089	0.234	-0.380	0.704
Spring 1st	EDAVG	0.592	0.900	0.658	0.511
Spring 1st	PRESCHY	7.524	5.810	1.295	0.197
Spring 1st	FEMALE	2.433	4.721	0.515	0.607
Spring 1st	OCAVG	3.342	1.449	2.307	0.022
Spring 1st	AGEMOTH	8.294	5.197	1.596	0.112
Spring 1st	BOOKSHM	6.786	1.915	3.543	0.000
Spring 1st	ANYEPTOP	3.819	6.603	0.578	0.564

Summary

A series of sensitivity analyses were completed to determine whether non-program-related factors might have influenced reading achievement. Adding covariates does not change the finding that school 30, the district

D school, and the schools in district G have significantly higher reading subtest scores at "spring first grade" than the schools in district E.

How were the K-1 analyses for reading affected by controlling for pretest?

The analyses of the reading achievement test scores in spring of first grade are presented in Table 246 through Table 249. Reducing to the students with pretest scores available, and therefore omitting district G, somewhat decreases the estimated difference between school 30 and the district E schools. The estimated difference of 22.903 falls just short of statistical significance (p = .077). Adding the pretest total as a predictor of spring first grade achievement further reduces the estimated SCHOOL30 effect to 17.281 (p = .096).

Table 246

Late-Exit K-1 Analysis for Reading
Pretest Students; Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	205.768	9.520	21.614	0.000
Spring 1st	OCAVG	2.268	1.833	1.237	0.217
Spring 1st	BOOKSHM	5.574	2.201	2.532	0.012

Table 247

Late-Exit K-1 Analysis for Reading
Pretest Students; School 30 and Key Covariates

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	208.433	8.421	24.752	0.000
Spring 1st	SCHOOL30	22.903	8.638	2.651	0.077
Spring 1st	OCAVG	1.374	1.817	0.756	0.451
Spring 1st	BOOKSHM	4.764	2.186	2.179	0.030

Table 248

Late-Exit K-1 Analysis for Reading
Pretest Students; Key Covariates and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	147.236	19.556	7.529	0.000
Spring 1st	OCAVG	2.195	1.743	1.259	0.209
Spring 1st	BOOKSHM	4.800	2.141	2.242	0.026
Spring 1st	ANALTOBE	0.176	0.054	3.284	0.001

Table 249

Late-Exit K-1 Analysis for Reading
Pretest Students; School 30, Key Covariates, and Pretest Total

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 1st	BASE	156.439	19.720	7.933	0.001
Spring 1st	SCHOOL30	17.281	8.451	2.045	0.096
Spring 1st	OCAVG	1.420	1.764	0.805	0.422
Spring 1st	BOOKSHM	4.208	2.132	1.973	0.050
Spring 1st	ANALTOBE	0.154	0.053	2.904	0.004

As was found for the language subtest, for the pretest students the parents' occupational status (OCAVG) is no longer a statistically significant predictor of the reading pretest score, but the number of books in the home is still a statistically significant predictor. The between-school variability is very small in the models that include SCHOOL30 as a predictor: the estimated between-school standard error is only 2.1 points in the model without the pretest and 1.3 in the model with the pretest. The within-school variability is reduced only to about 32 ESS points.

Summary

A sensitivity test was completed to determine whether excluding pretest information from the main K-1 reading analysis might have affected

ERIC

the results. Thus, the original K-1 HLM analysis of reading scores was done using only those students for whom there was a pretest score. If the results are similar, this would strengthen the original findings. Reducing the students to those with pretest scores available lowers the estimated difference between school 30 and the district E schools on reading achievement and the difference is no longer statistically significant. The between-school variation in district E is small.

What were the conclusions from the K-1 analyses for the reading subtest?

For the reading subtest, school 30 in district D and the schools in district G have significantly higher reading subtest scores at "spring first grade" than the schools in district E. The addition of covariates, including the pretest total, to the model reduces the estimated difference between school 30 and the district E schools from over 40 points to less than 20 points and it fails to achieve statistical significance in some analyses.

The considerable school-to-school variability on the reading subtest is nearly all explained by district differences. The estimated between-schools standard error is reduced to 6.35 points or less, compared to a within-school standard error of 32 points or more.

Results of the Late-Exit 1-6 Analyses

The purpose of the 1-6 HLM analysis was to describe the growth of these students from first grade through sixth grade. These analyses also allow us to describe specifically the students' growth from first grade to third grade and from third grade through sixth grade. For those who have hypotheses regarding the nature of the growth of late-exit students, these analyses provide an example of this growth.

ERIC

Why do the basic 1-6 analyses exclude the students in the K-3 cohort in district G?

The basic 1-6 analyses do not include the students in the K-3 cohort in district G. For these students, the "spring first grade" test scores were obtained in fall of second grade. Presumably these scores are higher than they would have been in spring of first grade, but the amount of the distortion is unknown. In an extensive series of analyses not presented here, models that accommodated this difference in test date were developed and estimated. These models were complex and did not add appreciably to the understanding of achievement growth in the late-exit districts.

Additional 1-6 models were developed using only the spring second grade and spring third grade test scores for the students in the K-3 cohort in district G. These models rely on a distinctively different data pattern for the schools in district G compared with the one school in district D and the schools in district E. The absence of a spring first grade data point in district G means that the estimated growth rate from first through third grade in that district is estimated using the growth from spring of second grade to spring of third grade.

The assumption that growth from second to third grade in district G is an unbiased estimate of growth from first to third grade is questionable. The fact that curvature from first to third grade is small in districts D and E indicates that for those two districts the assumption that growth is the same from first to second and from second to third grade may be reasonable. But even if it could be proven that the assumption of equal growth from first to second and second to third grades holds in districts D and E, that is not sufficient to ensure the assumption is tenable for district G where it is untestable. There are too many differences between district G and the other two late-exit districts to take that step.

For this reason, the basic 1-6 model was estimated without using the students from the K-3 cohort in district G. For comparison, the basic 1-6 model was estimated again after including the spring second grade and spring third grade test scores for the K-3 cohort in district G.

<u>What were the results of the basic 1-6 analysis for the mathematics subtest?</u>

The analysis of first grade through sixth grade mathematics achievement for the late-exit students requires combining students from two separate cohorts. The students in the K-3 cohort with first grade test scores and at least one more test score (second grade or third grade or both) provided information on growth from first through third grade. The students in the 3-6 cohort with at least two test scores provided information on growth from third through sixth grade. Both cohorts provided information on the level of achievement in third grade. As explained above, these basic 1-6 analyses do not include the K-3 cohort in district G because the "spring first grade" test scores were actually obtained in fall of second grade. The students in the 3-6 cohort in district G are included in the model.

The model of first through sixth grade achievement on the mathematics subtest is summarized in Table 250. The interpretation of this table and the corresponding tables for the language and reading subtests is slightly different from the interpretation of the tables giving results of the 1-3 analyses in Chapters III and IV. In this late-exit analysis, the intercepts of the individual growth curves are at spring of third grade; in the IS/EE 1-3 analyses in Chapters III and IV, the intercepts of the individual growth curves are at spring of first grade. This choice of intercept simplifies the comparison of the K-3 and 3-6 cohorts at spring of third grade. The lines labeled "Spring 3rd" in the "Predicted" column therefore represent the "initial status" for the 3-6 cohort but the "ending status" for the 1-3 cohort. The "Growth Rate" lines in the "Predicted" column represent the growth rate from grades 1-3 (for the K-3 cohort) or the growth rate from grades 3-6 (for the 3-6 cohort). There is no coefficient for curvature in these models, as explained above.

Table 250

Late-Exit 1-6 Analysis for Mathematics
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	296.913	12.964	22.904	0.000
Spring 3rd	EGRADE3	9.184	5.919	1.552	0.122
Spring 3rd	E3DISTG	-1.010	8.136	-0.124	0.901
Spring 3rd	EOSCHL44	-29.157	12.116	-2.407	0.017
Spring 3rd	SCHOOL45	-15.925	6.589	-2.417	0.016
Spring 3rd	AVABS	-0.880	0.234	-3.752	0.000
Spring 3rd	FEMALE	5.858	4.457	1.314	0.190
Spring 3rd	OCAVG	3.756	1.148	3.271	0.001
Spring 3rd	AGEMOTH	14.738	5.054	2.916	0.004
Spring 3rd	BOOKSHM	6.464	1.541	4.193	0.000
Growth Rate	BASE	34.735	6.215	5.589	0.000
Growth Rate	EGRADE3	-23.707	3.086	-7.682	0.000
Growth Rate	E3DISTG	-14.609	3.620	-4.035	0.000
Growth Rate	EOSCHL44	-18.665	6.590	-2.832	0.005
Growth Rate	FEMALE	6.561	2.347	2.795	0.006
Growth Rate	AGEMOTH	6.601	2.646	2.495	0.013

As in the other tables of results from hierarchical linear models, the "Predictor" column gives the names of variables used to predict the spring third grade test score or the growth rate. The "Parameter" column gives the estimated coefficient of the predictor in the "between-child" model and the "Std Err" column gives the estimated standard error of that parameter. The "t-Stat" and "p-Value" columns provide an approximate statistical test of whether the parameter is significantly different from zero.

The predictors of spring third grade mathematics achievement include an overall constant (BASE), four zero-one variables (EGRADE3, E3DISTG, EOSCHL44, and SCHOOL45), and five covariates (AVABS, FEMALE, OCAVG, AGEMOTH, and BOOKSHM). The four zero-one variables are used to accommodate differences among groups of children as explained below. The five covariates are included to adjust the estimated differences among the groups of students for any differences on these covariates and to improve the prediction of mathematics achievement. As usual, some nonsignificant covariates were included to facilitate comparisons with other models. AVABS is the average number of absences per year for the student, either between first and third grade (for the K-3 cohort) or between third and sixth grade (for the 3-6 cohort). FEMALE is 1 if the student is a girl and 0 if the student is a boy. OCAVG is the average of the parents' occupational status, AGEMOTH is the mother's age in three categories, and BOOKSHM is the number of books in the home. For more information on these covariates, see Chapter II.

The two covariates found to be predictive of mathematics growth rate were FEMALE and AGEMOTH. Both have positive coefficients, indicating that girls and those with older mothers have higher mathematics growth rates. The AGEMOTH variable is coded in three categories according to the age of the mother when the student entered the program in kindergarten: under 18, 18 to 25, and over 25 years old. Both FEMALE and AGEMOTH are also included as predictors of mathematics achievement at third grade. At spring of third grade, girls have slightly higher achievement than boys, but the difference is not statistically significant (difference = 5.858, $p = .190$). The difference between categories of mother's age at spring of third grade is statistically significant (difference = 14.738, $p = .004$).

For covariates included as predictors of mathematics achievement levels at spring of third grade but not as predictors of the growth rate, the estimated effect at spring of third grade is the same as the estimated effect at spring of every grade. The estimated effect does not change from one grade to another; it does not affect the rate of growth. That is, because the predicted growth rate is not affected by the covariate, the estimated difference between two students that is associated with the covariate is the same at every grade from first through sixth. For the two covariates FEMALE and AGEMOTH that are included as predictors of the growth rate, the estimated effect varies by grade.

For example, to find the estimated effect of FEMALE at spring of any grade other than third grade, the estimated difference in growth rate must

ERIC

be taken into account. The estimated difference between girls and boys at spring of third grade is 5.858, but the difference at spring of second grade is estimated to be 5.858 - 6.561 = -0.703 because he girls have a growth rate higher by 6.561 points per year. Similarly, the difference between boys and girls is estimated to be -7.264 in spring of first grade. Under this model, at the higher grades girls increase their lead over boys: the estimated difference is 5.858 + 6.561 = 12.419 at spring of fourth grade, 5.858 + 2 x 6.561 = 18.980 at spring of fifth grade, and 5.858 + 3 x 6.561 = 25.541 at spring of sixth grade. It should be noted that the estimated difference is more accurately estimated near the middle of the range of grades than it is at the lowest or highest grades.

It is important to avoid overinterpreting the coefficients for the covariates. No effort was made to test whether the effect of the covariates was different in the K-3 cohort than in the 3-6 cohort, or whether the effect of the covariates was different across the districts or schools. The purpose of including the covariates was not to evaluate the covariates themselves, but rather to evaluate district, school, and cohort differences after adjusting for the covariates.

Differences among districts, schools, and cohorts are represented in the model as zero-one variables used to predict the intercept ("Spring 3rd") or the slope ("Growth Rate") of the individual growth curves. Four zero-one variables are used as predictors of spring third grade mathematics achievement level. This allows different groups of children to have different levels of spring third grade test scores even after adjusting for the covariates. Each of the variables represents a comparison of one group of students with another, larger, group of students. The BASE coefficient gives the estimated mathematics achievement level for all students not explicitly represented by zero-one variables. EGRADE3 represents all students in the 3-6 cohort. E3DISTG represents the 3-6 cohort for district G. (The students in the K-3 cohort for district G are not included in the analysis, so E3DISTG represents all the students in district G who were included in the analysis; the variable is called E3DISTG instead of simply DISTG as a reminder that only the 3-6 cohort is included.) E0SCHL44 represents the K-3 cohort ("entry grade 0") for school 44 in district E. SCHOOL45 represents all the students for school 45 in district E. Thus the BASE coefficient is the estimated mathematics achievement level at spring of third grade for the K-3 students in the other schools: school 30 in district D, and schools 40, 41, 43, and 46 in district E. The E0SCHL44 and SCHOOL45 variables were included as predictors of "Spring 3rd" to allow the associated students to have third grade status different from the rest of the students. These groups of students were found to differ significantly from students in the other schools and cohorts at spring of third grade.

The variable EGRADE3 is included as a predictor of spring third grade mathematics achievement to allow the K-3 and 3-6 cohorts to have different levels at spring of third grade. EGRADE3 has the value 1 for students in the 3-6 cohort and the value 0 for students in the K-3 cohort. Thus the value of the parameter associated with EGRADE3 as a predictor of the status at spring of third grade is an estimate of how much higher the 3-6 cohort is than the K-3 cohort at that grade. When the parameter for EGRADE3 is positive, it indicates that the 3-6 cohort had higher achievement when tested in third grade (the first year of data collection) than the K-3 cohort had when tested in third grade (the fourth year of data collection). When the EGRADE3 parameter is not significantly different from zero, as is the case in Table 250, the two sets of third grade test scores are approximately the same (difference = 9.184, p = .122).

Because the K-3 district G students are not included in these basic 1-6 analyses, only the mathematics achievement among the students in the 3-6 cohort in district G can be compared with the other students. The E3DISTG variable as a predictor of spring of third grade provides an estimate of the difference between the 3-6 students in district G and the other 3-6 students at spring of third grade. The difference is very small and not statistically significant (difference = -1.010, p = .901). This means that the 3-6 cohort of students in district G schools have about the same level of mathematics achievement at spring of third grade as the 3-6 cohorts in the other schools.

In Table 250, the variable E0SCHL44 has a significantly negative coefficient as a predictor of mathematics test scores at spring of third grade (difference = -29.157, p = .017). This means that the students in the K-3 cohort in school 44 have a significantly lower level of mathematics achievement in spring of third grade than the students in the other K-3 cohorts. The 3-6 cohort in school 44 is not explicitly referenced in this model, indicating that the 3-6 cohort in school 44 does not have significantly different achievement at third grade compared with the 3-6 cohorts in the other schools. This implies that the K-3 and 3-6 cohorts in school 44 do not match well at spring of third grade. It should be noted that this conclusion and other findings about school cohorts should be interpreted cautiously. In school 44, there are only 20 students in the K-3 cohort and 19 in the 3-6 cohort (see Chapter II). With so few students, it is difficult to draw firm conclusions. A model that permitted both cohorts in school 44 to have lower achievement at spring of third grade fit the data nearly as well as the model presented in Table 250. It is debatable which model should be considered simpler; both models require the same number of parameters. Accordingly, the better-fitting model was selected for presentation. It should be emphasized that the difference in fit was small.

The negative coefficient for SCHOOL45 as a predictor of the spring third grade score indicates that both the K-3 and 3-6 cohorts in school 45 have relatively low mathematics achievement at spring of third grade (difference = -15.925, p = .016). The selection of this model with SCHOOL45 but neither E0SCHL45 nor E3SCHL45 for presentation implies that the "null hypothesis" of a single overall school difference could not be rejected. That is, the K-3 and 3-6 cohorts in school 45 have about the same relationship to each other as the K-3 and 3-6 cohorts do in the other schools. The EGRADE3 coefficient indicates that the 3-6 cohort is slightly higher (but not significantly higher) at third grade than the K-3 cohort; the hypothesis that this relationship holds in school 45 cannot be rejected.

Although there is no explicit curvature variable in the late-exit models, a form of curvature is accommodated by permitting the growth rate to be different in the two cohorts. This is accomplished by including the variable EGRADE3 as a predictor of the growth rate. The parameter associated with EGRADE3 as a predictor of growth rate is an estimate of the difference in growth rate between the K-3 cohort and the 3-6 cohort. The negative coefficient for EGRADE3 as a predictor of mathematics growth rate indicates that the growth rate is lower in grades 3-6 than in grades 1-3. This is consistent with the overall flattening of the national norms as shown in Chapter II. It is also consistent with the negative coefficients for curvature in the 1-3 models presented in Chapters III and IV. The annual increase in mathematics scores is estimated to be nearly 24 points lower from third to sixth grade than from first to third grade (difference = -23.707, p less than .0005). Of course, these two estimates are based on entirely separate groups of students. The growth rate from first to third grade is based on the K-3 cohort of students and the growth rate from third to sixth grade is based on the 3-6 cohort of students.

The students in the 3-6 cohort in district G schools have a lower estimated mathematics growth rate than the students in the 3-6 cohort in other schools (difference -14.609, p less than .0005). This implies an estimated growth rate for the 3-6 cohort in district G about 38.3 points per year lower than the K-3 cohorts in schools 30, 40, 41, 43, 45, and 46. The K-3 cohort in district G is not included in these analyses, so no comparisons with those students is possible here.

In school 44, the K-3 cohort is estimated to have a growth rate from first to third grade that is about 18.7 points lower than the other K-3 cohorts. This should be interpreted in connection with the estimated 29.2 point difference in spring of third grade. Taken together, these estimated differences in growth rate and status at third grade imply that school 44 students have mathematics achievement about 8.2 points higher than the typical students in the other schools (-29.157 + (-2) x (-18.665) = 8.180). This is less than the standard errors of the parameters, so it may be considered small. That is, the K-3 cohort in school 44 has about the same level of mathematics achievement in spring of first grade. However, by spring of third grade the school 44 students in the K-3 cohort are estimated to be about 29.2 points behind the students in other schools.

In contrast, students in the 3-6 cohort in school 44 do not have significantly different mathematics growth than the other schools in district E and the one school in district D. Thus the K-3 cohort in school 44 has lower mathematics scores in third grade, during the fourth year of data collection, than the 3-6 cohort in school 44 has in third grade, during the first year of data collection. It would be instructive to try to relate this difference to changes in the school or in the student population, but such analyses go beyond the scope of nominal program analyses.

Summary

The objective of the 1-6 analyses was to describe the growth in mathematics skills for students in the different late-exit districts. The 1-6 analysis for the mathematics subtest show a pattern of student growth that is fairly consistent, but there are some district and school differences. The K-3 cohort of students exhibits significantly higher growth in first through third grade than the 3-6 cohort exhibits in third through sixth grade. This is consistent with the flattening growth indicated by the national norms. The only district difference found is that the 3-6 cohort in district G has lower growth than the other 3-6 cohorts; the K-3 cohort in district G was not included in the analysis and so cannot be compared directly. It is noteworthy that the 3-6 cohorts in district G schools have almost exactly the same status at third grade as the 3-6 cohorts in other schools. School 30, the only school in district D, was not found to differ significantly from the other schools in mathematics achievement. School differences were found for schools 44 and

ERIC

45 in district E. Although the mathematics growth in school 45 from first through sixth grades parallels the growth for other schools in district E, the students in school 45 have consistently lower mathematics subtest scores by about 16 ESS points. The pattern of mathematics growth in the K-3 cohort in school 44 is unusual. The K-3 cohort of students in school 44 have about the same level of spring first grade mathematics achievement as in other schools but the growth is substantially lower, resulting in scores about 29 points lower at spring of third grade. The 3-6 cohort of students in school 44 have growth that is similar to that among the other 3-6 cohorts in district E.

What were the results of the basic 1-6 analysis for the language subtest?

The objective of the basic analysis for English language skills is to describe the growth in this area among students in the different late-exit districts. Table 251 summarizes the results of the 1-6 analysis for the language subtest. Compared to the analysis for the mathematics subtest (see Table 250), there is only one change among the variables used to predict the spring third grade status and no changes among the predictors of the growth rate. The variable SCHOOL30 replaces the variable SCHOOL45 as a predictor of spring third grade status. For the language subtest, the HLM analysis required a separate intercept for SCHOOL30 as SCHOOL30 was found to be significantly different from the other schools. In the HLM analysis of the mathematics subtest, it was SCHOOL45 that required a separate intercept.

Table 251

Late-Exit 1-6 Analysis for Language
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	275.788	15.543	17.744	0.000
Spring 3rd	EGRADE3	6.320	6.871	0.920	0.359
Spring 3rd	E3DISTG	2.380	9.575	0.249	0.804
Spring 3rd	EOSCHL44	-65.041	13.785	-4.718	0.000
Spring 3rd	SCHOOL30	50.400	7.808	6.455	0.000
Spring 3rd	AVABS	-0.743	0.296	-2.506	0.013
Spring 3rd	FEMALE	15.838	5.266	3.007	0.003
Spring 3rd	OCAVG	3.389	1.610	2.105	0.036
Spring 3rd	AGEMOTH	17.728	5.961	2.974	0.003
Spring 3rd	BOOKSHM	10.356	2.092	4.951	0.000
Growth Rate	BASE	43.328	6.272	6.909	0.000
Growth Rate	EGRADE3	-32.328	3.140	-10.297	0.000
Growth Rate	E3DISTG	-0.832	3.567	-0.233	0.816
Growth Rate	EOSCHL44	-43.648	6.789	-6.429	0.000
Growth Rate	FEMALE	5.191	2.372	2.189	0.029
Growth Rate	AGEMOTH	5.303	2.656	1.997	0.047

For the language subtest, the estimated coefficients for the covariates are similar in magnitude and are of the same sign as for the mathematics subtest. One difference between the two subtests appears in the relationship between the achievement for boys and girls. For the mathematics subtest, girls were estimated to be slightly behind boys at spring of first grade, about even in second grade, and increasingly ahead from third grade through sixth grade. For the language subtest, girls are estimated to be ahead by about 5.5 points and to increase that advantage by about 5.2 points per year.

As was found for the mathematics subtest, for the language subtest the growth rate from third through sixth grade (for the 3-6 cohort) is significantly lower than the growth rate from first through third grade (for the 1-3 cohort). Again, this is consistent with the pattern of growth indicated by the national norms. The third grade scores in the K-3 and 3-6 cohorts match well in all schools except school 44 in district E,

ERIC

as was found for the mathematics subtest. For the language subtest, the students in the 3-6 cohort in district G exhibited growth rates from third through sixth grades that were virtually identical to those of the students in 3-6 cohorts in district E (difference = -0.832, p = .804). In contrast, the 3-6 cohort in district G had similar third grade status but markedly lower growth from third through sixth grades on the mathematics subtest.

For the mathematics subtest, students in school 45 were found to have significantly lower test scores but the same pattern of growth. For the language subtest, no significant difference was found between the students in school 45 and the students in the other schools. However, the students in school 30, the only school in district D, had language test scores that were about 50 points higher than those for the students in districts E and G. It should be emphasized that the comparison with district G students extends only to the 3-6 cohort in district G, since the K-3 cohort of students in district G is not included in this analysis. These higher achievement levels on the language subtest in school 30 are in contrast to the finding of no significant differences between school 30 and the district E schools for the mathematics subtest.

The students in the K-3 cohort of school 44 exhibit an unusual pattern of growth on the language subtest, just as they did for the mathematics subtest. These students have an estimated advantage at spring of first grade of about 22.3 points compared with the students in other district E schools. With estimated growth from first to third grade about 43.6 points lower than in the other district E schools, at spring of third grade the K-3 cohort in school 44 is about 65 points lower. As was found for the mathematics subtest, the 3-6 cohort of students in school 44 was not significantly different from the other 3-6 cohorts.

Summary

The purpose of the 1-6 analyses of English language skills was to describe the academic growth of students in late-exit districts from first grade to sixth grade. The 1-6 analysis of the language subtest showed somewhat different school and district differences than the mathematics subtest. For both the mathematics and language subtests, the growth rate from third through sixth grades (for the 3-6 cohort) is significantly lower than the growth rate from first through third grades (for the K-3 cohort). The 3-6 cohort of students in district G schools exhibited approximately the same levels of language achievement as the 3-6 cohorts in other schools, in contrast to the finding of significantly lower mathematics growth rates for the 3-6 students in district G. (The K-3 cohort of students in district G is not included in this set of analyses.) The school 45 students had approximately the same levels of language achievement as the students in the other district E schools, unlike the finding of lower mathematics achievement for school 45. For the language subtest, the students in school 30, the only study school in district D, had achievement scores about 50 points higher than the students in other schools. For the mathematics subtest, school 30 was not found to differ from the schools in district E. The students in the K-3 cohort in school 44 were found to have significantly lower language growth rates from first grade to third grade despite having first grade status about as high or higher than the students in other schools. This is the same pattern that was found for the mathematics subtest.

What were the results of the basic 1-6 analysis for the reading subtest?

Table 252 summarizes the results of the 1-6 analysis for the reading subtest. The same variables are used to predict the status at spring third grade and the growth rate as were used for the language subtest (see Table 251). The results are remarkably similar for the reading and language subtests. The main difference between the two analyses is that the effect of many of the predictor variables is weaker for the reading subtest. Among the covariates, several are no longer statistically significant predictors or are only marginally significant. Only the average number of absences (AVABS) and the number of books in the home (BOOKSHM) are statistically significant at the .01 level in the presence of the other covariates.

Table 252

Late-Exit 1-6 Analysis for Reading
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	288.457	15.241	18.927	0.000
Spring 3rd	EGRADE3	3.807	6.810	0.559	0.577
Spring 3rd	E3DISTG	-13.226	9.475	-1.396	0.164
Spring 3rd	EOSCHL44	-49.279	13.738	-3.587	0.000
Spring 3rd	SCHOOL30	35.006	7.609	4.601	0.000
Spring 3rd	AVABS	-0.790	0.290	-2.724	0.007
Spring 3rd	FEMALE	8.665	5.158	1.680	0.094
Spring 3rd	OCAVG	2.615	1.569	1.666	0.097
Spring 3rd	AGEMOTH	11.672	5.841	1.998	0.047
Spring 3rd	BOOKSHM	9.136	2.044	4.470	0.000
Growth Rate	BASE	52.594	6.358	8.272	0.000
Growth Rate	EGRADE3	-27.995	3.211	-8.719	0.000
Growth Rate	E3DISTG	-1.764	3.641	-0.484	0.628
Growth Rate	EOSCHL44	-32.144	6.949	-4.625	0.000
Growth Rate	FEMALE	3.247	2.398	1.354	0.177
Growth Rate	AGEMOTH	2.064	2.688	0.768	0.443

As was found for both the mathematics subtest and the language subtest, for the reading subtest the growth rate from third through sixth grades (for the 3-6 cohort) is significantly lower than the growth rate from first through third grades (for the 1-3 cohort). The national norm information makes this as unsurprising for the reading subtest as it was for the other two subtests.

The 3-6 cohort of students in district G show a pattern of reading achievement that is similar to that in the other 3-6 cohorts, just as was found for the language subtest. The 3-6 cohort in district G does have slightly lower estimated reading test scores at spring of third grade, but

ERIC

the difference is not statistically significant (difference = -13.226, p = .164).

Just as for the language subtest, the students in school 30, the only school in district D, had reading test scores significantly higher than those for the students in other schools (difference = 35.006, p less than .0005).

The students in the K-3 cohort of school 44 exhibit the same unusual pattern of growth on the reading subtest as they did on the other two subtests. These students have an estimated advantage in reading achievement at spring of first grade of about 15.0 points, but with a growth rate about 32.1 points lower they have an estimated disadvantage of 49.3 points at spring of third grade. Again, the 3-6 cohort of students in school 44 showed no significant differences in reading achievement compared with the 3-6 cohorts in other schools.

<u>Summary</u>

The primary objective of the 1-6 analysis of reading achievement was to describe the growth in reading skills among students in the different late-exit districts. The 1-6 analysis of the reading subtest showed essentially the same results as the corresponding analysis for the language subtest. For both subtests, the growth rate from third through sixth grades (for the 3-6 cohort) is significantly lower than the growth rate from first through third grades (for the K-3 cohort). The 3-6 cohort of students in district G schools exhibited approximately the same levels of reading and language achievement as the 3-6 cohorts in other schools. For both subtests, the students in district D (school 30) had achievement scores significantly higher than the students in other schools. The students in the K-3 cohort in school 44 had significantly lower growth rates in both reading and language from first grade to third grade despite having first grade status about as high as or higher than the students in other schools.

How were the results of the 1-6 analyses affected by considering only the K-3 cohort?

As a test of the sensitivity of the 1-6 analyses, for each of the subtests the basic 1-6 model was estimated using only the students from the K-3 cohort. For these students, a 1-3 model was estimated using the same predictor variables as were used in the 1-6 analyses except for EGRADE3 and E3DISTG, which apply only to the 3-6 cohort.

Table 253 summarizes the results of the 1-3 analysis for the mathematics subtest; it should be compared with Table 250. The results of the 1-3 analysis for mathematics are very similar to the results of the 1-6 analysis. Almost all of the parameter estimates in the 1-3 analysis are within one standard error of the corresponding estimate for the 1-6 analysis. Although some covariates are no longer statistically significant, the school differences all remain statistically significant. No alteration in the conclusions for the mathematics subtest is warranted.

Table 254 and Table 255 show the results of the 1-3 analyses for the language and reading subtests, respectively. These are even closer to the corresponding models estimated in the 1-6 analyses, Table 251 and Table 252. There are no appreciable differences in the school effects for either the language subtest or the reading subtest.

Table 253

Late-Exit 1-3 Analysis for Mathematics
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	314.428	19.694	15.966	0.000
Spring 3rd	EOSCHL44	-35.367	12.006	-2.946	0.004
Spring 3rd	SCHOOL45	-18.152	7.094	-2.559	0.012
Spring 3rd	AVABS	-0.521	0.309	-1.687	0.094
Spring 3rd	FEMALE	7.834	8.347	0.939	0.350
Spring 3rd	OCAVG	2.944	1.233	2.387	0.019
Spring 3rd	AGEMOTH	9.791	8.345	1.173	0.243
Spring 3rd	BOOKSHM	4.077	1.781	2.289	0.024
Growth Rate 1-3	BASE	39.039	9.110	4.285	0.000
Growth Rate 1-3	EOSCHL44	-18.114	5.770	-3.139	0.002
Growth Rate 1-3	FEMALE	8.807	4.122	2.136	0.035
Growth Rate 1-3	AGEMOTH	4.226	4.163	1.015	0.312

Table 254

Late-Exit 1-3 Analysis for Language
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	280.050	28.541	9.812	0.000
Spring 3rd	EOSCHL44	-72.420	17.226	-4.204	0.000
Spring 3rd	SCHOOL30	54.987	9.533	5.768	0.000
Spring 3rd	AVABS	-0.246	0.466	-0.528	0.598
Spring 3rd	FEMALE	17.114	12.004	1.426	0.157
Spring 3rd	OCAVG	2.832	1.965	1.442	0.152
Spring 3rd	AGEMOTH	15.974	12.015	1.329	0.186
Spring 3rd	BOOKSHM	9.313	2.879	3.235	0.002
Growth Rate 1-3	BASE	42.542	13.149	3.235	0.002
Growth Rate 1-3	EOSCHL44	-44.297	8.352	-5.304	0.000
Growth Rate 1-3	FEMALE	4.783	5.950	0.804	0.423
Growth Rate 1-3	AGEMOTH	6.148	6.017	1.022	0.309

Table 255

Late-Exit 1-3 Analysis for Reading
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	298.014	24.884	11.976	0.000
Spring 3rd	EOSCHL44	-59.667	14.878	-4.010	0.000
Spring 3rd	SCHOOL30	39.507	8.442	4.680	0.000
Spring 3rd	AVABS	0.022	0.412	0.054	0.957
Spring 3rd	FEMALE	10.055	10.454	0.962	0.338
Spring 3rd	OCAVG	2.466	1.739	1.419	0.159
Spring 3rd	AGEMOTH	6.240	10.434	0.598	0.551
Spring 3rd	BOOKSHM	6.884	2.548	2.702	0.008
Growth Rate 1-3	BASE	59.155	10.616	5.572	0.000
Growth Rate 1-3	EOSCHL44	-32.724	6.734	-4.859	0.000
Growth Rate 1-3	FEMALE	3.078	4.804	0.641	0.523
Growth Rate 1-3	AGEMOTH	-0.743	4.854	-0.153	0.879

For all three subtests, perhaps the most noteworthy change resulting from the reduction to the K-3 cohort involves the average number of absences. In the 1-6 models, the effect of the average number of absences was fairly similar for all three subtests (-0.880, -0.743, and -0.790) and consistently significant (p = .013 or less). In the corresponding models including only first through third grades, the coefficient of AVABS is considerably reduced (-0.521, -0.246, and 0.022) and is never statistically significant (p = .094 or higher). This indicates that the number of days absent may be a less important predictor of achievement in the early primary grades than it is in the later primary grades.

Summary

As a test of the sensitivity of the 1-6 analyses for each subtest (i.e., mathematics, English language, and English reading), the basic 1-6 model was estimated using the same predictor variables as were used in the 1-6 analyses. The results of the 1-3 analyses are very similar to the results of the basic 1-6 analyses. In other words, including only the students in the K-3 cohort does not materially change the estimated models.

How were the results of the 1-6 analyses affected by considering only the 3-6 cohort?

As another test of the sensitivity of the 1-6 analyses, a 3-6 model was estimated using the students in the 3-6 cohort. The same predictor variables were used in the 3-6 model as were used in the 1-6 model with two exceptions: EGRADE3 and EOSCHL44. The variable EGRADE3 has the value 1 for all the students in the 3-6 cohort, so it would be the same as the BASE variable. The EOSCHL44 indicates the K-3 cohort in school 44 and therefore cannot be included in the analyses of the 3-6 cohort.

Table 256 through Table 258 summarize the results of the 3-6 analyses for the mathematics, language, and reading subtests. The results of the 3-6 analyses for all three subtests are very similar to the results of the corresponding 1-6 analyses. Almost every parameter estimate in the 3-6 analyses is within one standard error of the corresponding estimate in the 1-6 analyses. Note that the BASE coefficients in the 3-6 analyses need to be compared to the sum of the BASE and the EGRADE3 coefficients in the 1-6 analyses. The BASE for spring of third grade in the 3-6 analyses is consistently lower than the sum of BASE and EGRADE3 as predictors of spring third grade in the 1-6 analyses. However, the change is always within 1.5 times the standard error in the 3-6 analysis and therefore should be considered fairly small.

Table 256

Late-Exit 3-6 Analysis for Mathematics
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	275.458	20.886	13.189	0.000
Spring 3rd	E3DISTG	2.176	8.967	0.243	0.809
Spring 3rd	SCHOOL45	-11.381	13.912	-0.818	0.415
Spring 3rd	AVABS	-1.045	0.341	-3.065	0.003
Spring 3rd	FEMALE	11.640	6.981	1.667	0.098
Spring 3rd	OCAVG	4.369	2.373	1.841	0.068
Spring 3rd	AGEMOTH	22.640	8.643	2.620	0.010
Spring 3rd	BOOKSHM	10.209	2.705	3.775	0.000
Growth Rate 3-6	BASE	16.814	9.644	1.743	0.083
Growth Rate 3-6	E3DISTG	-14.220	3.968	-3.584	0.000
Growth Rate 3-6	FEMALE	3.641	3.514	1.036	0.302
Growth Rate 3-6	AGEMOTH	4.790	4.113	1.165	0.246

Table 257

Late-Exit 3-6 Analysis for Language
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	258.049	24.730	10.435	0.000
Spring 3rd	E3DISTG	-0.526	10.181	-0.052	0.959
Spring 3rd	SCHOOL30	34.781	14.111	2.465	0.015
Spring 3rd	AVABS	-1.015	0.385	-2.635	0.009
Spring 3rd	FEMALE	15.067	7.822	1.926	0.056
Spring 3rd	OCAVG	3.125	2.856	1.094	0.276
Spring 3rd	AGEMOTH	31.480	10.061	3.129	0.002
Spring 3rd	BOOKSHM	10.488	3.072	3.414	0.001
Growth Rate 3-6	BASE	16.477	7.447	2.213	0.029
Growth Rate 3-6	E3DISTG	-0.732	3.052	-0.240	0.811
Growth Rate 3-6	FEMALE	5.074	2.714	1.870	0.064
Growth Rate 3-6	AGEMOTH	2.998	3.163	0.948	0.345

Table 258

Late-Exit 3-6 Analysis for Reading
Excluding K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	267.508	26.727	10.009	0.000
Spring 3rd	E3DISTG	-14.440	11.088	-1.302	0.195
Spring 3rd	SCHOOL30	19.379	14.669	1.321	0.189
Spring 3rd	AVABS	-1.092	0.404	-2.702	0.008
Spring 3rd	FEMALE	9.836	8.551	1.150	0.252
Spring 3rd	OCAVG	2.448	2.972	0.824	0.411
Spring 3rd	AGEMOTH	22.765	10.930	2.083	0.039
Spring 3rd	BOOKSHM	11.964	3.210	3.727	0.000
Growth Rate 3-6	BASE	26.289	9.004	2.920	0.004
Growth Rate 3-6	E3DISTG	-2.011	3.685	-0.546	0.586
Growth Rate 3-6	FEMALE	2.325	3.276	0.710	0.479
Growth Rate 3-6	AGEMOTH	1.811	3.828	0.473	0.637

Some covariates are no longer statistically significant, but all have the same sign and most have the same magnitude. As expected from the comparison of the 1-6 and 1-3 models, the average number of absences (AVABS) has a slightly larger effect in the 3-6 analyses than in the 1-6 analyses. The effect is similar for all three subtests (-1.045, -1.015, and -1.092), and it is always significant at the .01 level despite the smaller sample size. Another consistent change from the 1-6 models to the 3-6 models is that the effect of AGEMOTH as a predictor of spring third grade is always somewhat larger in the 3-6 models. Also, the coefficient of BOOKSHM as a predictor of spring third grade in the 3-6 models is a little larger for mathematics and reading although about the same for the language subtest.

The school effects are reduced in the 3-6 analyses compared with the 1-6 analyses, and none are significant at the .01 level. Only the coefficient for SCHOOL30 in the model for the language subtest achieves significance at the .05 level (difference = 34.781, p = .015). The reduced statistical significance is largely attributable to the increase in the standard error which in turn is largely attributable to the decrease in the number of students included in the 3-6 analysis.

Summary

As an additional sensitivity test of the 1-6 analysis, a 3-6 model was estimated using students from the 3-6 cohort. The results of the 3-6 analyses show more deviation from the basic 1-6 analyses than the 1-3 analyses showed, but the 3-6 analyses are still very similar to the 1-6 analyses. Including only the students in the 3-6 cohort reduces the estimated school differences, and the increase in the associated standard errors makes none of the school differences statistically significant at the .01 level. The school differences are still in the same direction: students in school 45 in district E exhibit relatively low mathematics growth rates, while students in school 30 in district D exhibit relatively high language and reading growth rates.

How were the results of the 1-6 analyses affected by including the K-3 cohort in district G?

As explained above, the basic 1-6 analyses exclude the students in the K-3 cohort in district G. They were excluded because the "spring first grade" test scores for these student are actually from fall of second grade. For purposes of comparison with the basic 1-6 analyses, the same models were recalculated after including the spring second grade and spring third grade test scores for the K-3 students in district G. This is reasonable under the untestable assumption that growth in district G is the same from spring of first grade to spring of second grade as it is from spring of second grade to spring of third grade. In support of this assumption for the students in district G is the fact that it holds approximately for the students in the other two late-exit districts: the amount of curvature in the 1-3 analyses for districts D and E is fairly small. It must be emphasized that there are so many differences between district G and the other two late-exit districts that it is not reasonable to rely on an assumption that district G students would exhibit the same pattern of growth as the other late-exit students.

Table 259 through Table 261 summarize the results of the 1-6 analyses that include the K-3 cohort in district G. These tables may be compared with Table 250 through Table 252. Most of the estimated parameters are similar. For the covariates and the school variables, the parameters are very close and the statistical significance is essentially unaffected. However, the addition of the 47 students with spring second grade and spring third grade scores in district G does make some material differences.

Table 259

Late-Exit 1-6 Analysis for Mathematics
Including K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	295.287	11.786	25.054	0.000
Spring 3rd	EGRADE3	11.601	5.054	2.295	0.022
Spring 3rd	DISTG	-5.559	5.041	-1.103	0.271
Spring 3rd	EOSCHL44	-27.760	11.798	-2.353	0.019
Spring 3rd	SCHOOL45	-15.936	6.479	-2.460	0.014
Spring 3rd	AVABS	-0.863	0.221	-3.901	0.000
Spring 3rd	FEMALE	5.813	4.058	1.432	0.153
Spring 3rd	OCAVG	3.679	1.041	3.534	0.000
Spring 3rd	AGEMOTH	15.256	4.640	3.288	0.001
Spring 3rd	BOOKSHM	6.040	1.397	4.323	0.000
Growth Rate	BASE	33.280	5.956	5.588	0.000
Growth Rate	EGRADE3	-22.563	2.776	-8.127	0.000
Growth Rate	DISTG	-15.086	2.934	-5.142	0.000
Growth Rate	EOSCHL44	-17.629	6.534	-2.698	0.007
Growth Rate	FEMALE	7.200	2.287	3.148	0.002
Growth Rate	AGEMOTH	6.646	2.574	2.582	0.010

Table 260

Late-Exit 1-6 Analysis for Language
Including K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	271.955	14.834	18.333	0.000
Spring 3rd	EGRADE3	4.797	6.164	0.778	0.437
Spring 3rd	DISTG	5.357	6.511	0.823	0.411
Spring 3rd	EOSCHL44	-66.159	14.029	-4.716	0.000
Spring 3rd	SCHOOL30	45.692	7.928	5.764	0.000
Spring 3rd	AVABS	-0.801	0.297	-2.695	0.007
Spring 3rd	FEMALE	14.641	5.040	2.905	0.004
Spring 3rd	OCAVG	4.707	1.534	3.068	0.002
Spring 3rd	AGEMOTH	17.809	5.752	3.096	0.002
Spring 3rd	BOOKSHM	11.264	1.981	5.686	0.000
Growth Rate	BASE	43.947	6.161	7.133	0.000
Growth Rate	EGRADE3	-32.947	2.896	-11.376	0.000
Growth Rate	DISTG	-0.756	2.948	-0.256	0.798
Growth Rate	EOSCHL44	-44.220	6.887	-6.421	0.000
Growth Rate	FEMALE	5.661	2.361	2.397	0.017
Growth Rate	AGEMOTH	5.187	2.644	1.962	0.051

Table 261

Late-Exit 1-6 Analysis for Reading
Including K-3 Students in District G

Predicted	Predictor	Parameter	Std Err	t-Stat	p-Value
Spring 3rd	BASE	289.869	13.808	20.993	0.000
Spring 3rd	EGRADE3	-0.306	5.810	-0.053	0.958
Spring 3rd	DISTG	1.486	6.061	0.245	0.806
Spring 3rd	EOSCHL44	-50.431	13.331	-3.783	0.000
Spring 3rd	SCHOOL30	31.225	7.217	4.327	0.000
Spring 3rd	AVABS	-0.871	0.275	-3.171	0.002
Spring 3rd	FEMALE	6.237	4.694	1.329	0.185
Spring 3rd	OCAVG	3.331	1.400	2.379	0.018
Spring 3rd	AGEMOTH	10.926	5.365	2.037	0.043
Spring 3rd	BOOKSHM	9.902	1.820	5.442	0.000
Growth Rate	BASE	51.402	6.069	8.469	0.000
Growth Rate	EGRADE3	-26.453	2.865	-9.233	0.000
Growth Rate	DISTG	-6.458	2.925	-2.208	0.028
Growth Rate	EOSCHL44	-31.606	6.811	-4.641	0.000
Growth Rate	FEMALE	3.832	2.323	1.650	0.100
Growth Rate	AGEMOTH	2.284	2.605	0.877	0.381

For the mathematics subtest, there are several differences. The estimated coefficient for EGRADE3 as a predictor of spring third grade is increased from 9.184 to 11.601 and is statistically significant (p = .022). This implies that the K-3 and 3-6 cohorts do not match well at spring of third grade. In addition, the coefficient of DISTG at spring of third grade is decreased from -1.010 to -5.559, although it is still not statistically significant (p = .271). However, the coefficients for EGRADE3 and DISTG as predictors of the growth rate are little changed.

For the language subtest, there are no dramatic differences, but for the reading subtest there are two. The estimated coefficient of DISTG as a predictor of the growth rate decreases from -1.764 (p = .628) to -6.458 (p = .028). This implies that the district G students have a significantly lower reading growth rate than the students in the other districts after including the additional 47 students. Also, the coefficient of DISTG as a predictor of spring third grade changes from -13.226 to 1.486,

ERIC

although neither value is significantly different from zero (p greater than .10).

Summary

As a sensitivity test of the basic 1-6 analyses which excluded district G because of different test dates, the basic analyses were rerun including the available district G data. The addition of the 47 students in district G with spring second grade and spring third grade scores results in some material changes for the mathematics and reading subtests; there are no material changes for the language subtest. For the mathematics subtest, the students in the K-3 and 3-6 cohorts across the three late-exit districts no longer match well in third grade. For the reading subtest, the growth rates for the students in district G are no longer similar to the growth rates for the students in the other two late-exit districts. Although it would be possible to investigate these changes in more detail to try to determine what aspect of the 47 pairs of scores produces these changes, the most important conclusion is clear: the basic 1-6 analyses are sensitive to the inclusion of the K-3 cohort of students in district G. Because the interpretation of the results of the model that includes the K- students in district G relies on the untestable assumption of steady growth from spring first grade to spring third grade in that district, no further analyses of this sensitivity were performed. The 1-6 analyses which exclude the K-3 students in district G do not rely on this untestable assumption and stand as the basic models for the late-exit districts.

How much of the student variability in achievement is explained by the 1-6 models?

In the 1-6 analyses, a hierarchical linear model is used to predict the intercept and slope of growth curves calculated for each student. The prediction models presented here include five covariates as predictors of the intercept (status at spring of third grade) plus several zero-one variables for comparing groups of students. Similarly, there are two covariates as predictors of the growth rate (either the 1-3 growth rate or the 3-6 growth rate, depending on which cohort the student is in) together with a similar set of zero-one variables. It is natural to ask how well these prediction models account for the student-to-student differences in achievement.

For all three subtests, there is considerable student variability in the estimated status at third grade that is not accounted for by the prediction model. For the mathematics subtest, the estimated student variability in third grade status is a standard error of about 30.8 points; for the language subtest, the standard error is about 38.0 points; and for the reading subtest, the standard error is about 36.8 points. All of these values indicate highly significant student-to-student variation in achievement levels not explained by the model. This variation is to be expected: there is a natural distribution of student ability that is unlikely to be captured by the background variables used as covariates.

In contrast, for all three subtests the student variability in growth is largely explained by the prediction model. For the mathematics subtest, the estimated between-student variability in mathematics growth rates is a standard error of 9.3 points; for the language subtest, the standard error is about 6.1 points; and for the reading subtest the standard error is also about 6.1 points. An approximate test that the remaining variability in growth rates is zero is nonsignificant for the mathematics and reading subtests ($p = .364$ and $p = .208$), and borderline for the language subtest ($p = .049$). In other words, the prediction model for the growth rates explains most of the student-to-student differences in those growth rates.

Conclusions

What are the overall conclusions from the late-exit analyses?

The objective of the late-exit analyses was to describe the growth in achievement by language-minority children from first grade through sixth grade. The K-1 analyses showed significant differences among the districts. School 30, the only school in district D, showed significantly higher language and reading test scores at spring of first grade than the schools in district E. The differences are reduced by including covariates, but remain statistically significant or at least borderline. Although school 30 showed significantly higher mathematics test scores at spring of first grade compared with the schools in district E, the difference is not significant once covariates are included in the model. The "spring first grade" tests in district G were administered in fall of second grade, so the higher "spring first grade" scores on the mathematics and reading subtests in district G compared with district E are expected. More surprising is that although the language scores in district G are somewhat higher than in district E, the difference is not statistically significant. That is, the district G students tested in fall of second grade had about the same level of language achievement as the district E students tested in spring of first grade.

The curvature within the K-3 cohort (from first grade to third grade) and within the 3-6 cohort (from third grade to sixth grade) was found to be small. Where the curvature was statistically significant, it was positive, indicating increasing growth. The 1-6 analyses presented used straight-line individual growth curves to avoid unnecessary complexity based on tenuous data.

The 1-6 analyses showed a consistent difference in growth rate between the early primary grades and the later primary grades. The growth rates from first to third grade, estimated using the students in the K-3 cohort, were found to be much higher than the growth rates from third to sixth grade, estimated using the students in the 3-6 cohort. This finding is in accordance with the flattening growth indicated by the national norms.

The relatively high level of language and reading achievement in the district D school, school 30, that was found in the K-1 analyses was confirmed by the 1-6 analyses. The 1-6 analyses also confirmed that the level of mathematics achievement was similar in school 30 and the district E schools after adjusting for covariates. The growth rates of the students in school 30 were not found to differ significantly from the growth rates in the district E schools on any of the three subtests.

The 1-6 analyses did uncover two school effects within district E. School 45 had significantly lower mathematics achievement than the other district E schools, but comparable growth rates. For all three subtests, the students in the K-3 cohort in school 45 had significantly lower growth rates from spring first grade to spring third grade and therefore significantly lower third grade scores despite comparable first grade scores. In contrast, the students in the 3-6 cohort in school 45 did not exhibit any significant deviations from the students in the 3-6 cohort in the other district E schools.

The 1-6 analyses were sensitive to the inclusion of the spring second grade and spring third grade scores for the students in the K-3 cohort in district G. Interpretation of the model with these test scores included requires making untestable assumptions, so the basic 1-6 analyses exclude the K-3 cohort in district G. The students in the 3-6 cohort in district G exhibited significantly lower growth rates on the mathematics subtest than the students in the 3-6 cohorts in district E. However, the language and reading growth rates for the students in the 3-6 cohorts in districts E and G were comparable.

The 1-6 models provided a reasonable fit to the data. With the exceptions noted above, the K-3 and 3-6 cohorts matched well at third grade. Most of the student variation in growth rates is explained by the prediction model. The growth models for spring first grade to spring

ERIC

third grade estimated using only the K-3 cohort were similar to the 1-6 models, as were the growth models for spring third grade to spring sixth grade estimated using only the 3-6 cohort.

ERIC

VI. COMPARISON OF PROGRAMS, DISTRICTS, AND THE NORMING POPULATION USING TRAJECTORY ANALYSIS OF MATCHED PERCENTILES (TAMP)

Introduction

How is this chapter related to the analyses of Chapters III, IV, and V?

This chapter serves as a supplement to the analyses in Chapters III, IV, and V. Those chapters presented the results of analyses designed to answer specific analytic questions. This chapter provides supplemental information about the achievement of the students in this study in the form of graphs. These graphs should not be considered a replacement for the analyses of the earlier chapters. The analyses in the earlier chapters are carefully controlled for differences among students, schools, and districts; they provide the most complete answers to the specific analytic questions. The graphs in this chapter do not reflect any adjustment for differences among students, schools, and districts and are therefore potentially misleading. Any conflict between the findings from the analyses of the earlier chapters and the graphs of this chapter must be resolved in favor of those carefully controlled analyses, not the uncontrolled graphs. For convenience, the conclusions from the three analytic chapters will be summarized here.

The analyses reported in Chapter III were designed to evaluate the relative effectiveness of immersion strategy and early-exit programs in schools with both programs. Only small differences were found between the immersion strategy program and the early-exit program in the four two-program schools. No statistically significant differences were found for the mathematics subtest. For both the language subtest and the reading subtest, there was no statistically significant program difference at spring of third grade. However, there was a temporary advantage for immersion strategy students at spring of first grade (for language) or spring of second grade (for reading).

The analyses reported in Chapter IV were designed to evaluate the relative effectiveness of immersion strategy and early-exit programs in schools with only one of the programs. The major finding was that any program differences are small compared to the school differences. No statistically significant differences were found for the mathematics subtest or the reading subtest. For the language subtest, there was no statistically significant program difference at spring of third grade and no statistically significant program difference in average growth rate from spring first grade to spring third grade. However, there was a temporary advantage for immersion strategy students at spring of second grade on the language subtest.

The analyses reported in Chapter V were designed to compare the patterns of growth for students in the late-exit program across the districts, schools, and cohorts. Substantial differences were found. The growth rate from spring first grade to spring third grade was found to be substantially greater than for spring third grade to spring sixth grade, in accordance with the flattening growth implied by the national norms. School 30, the only school in district D, had relatively high levels of language and reading achievement (but similar levels of mathematics achievement) compared with the schools in district E. The growth rates for students in school 30 were similar to the growth rates for students in the district E schools for all three subtests. As explained in Chapter V, students in the K-3 cohort in district G were not included in the basic model because the spring first grade test scores were actually obtained in fall second grade and the model was sensitive to their inclusion. For the 3-6 cohort in district G, the language and reading growth rates were similar to those for the 3-6 cohort in district E, but the mathematics growth rate was much lower.

<u>What is a Trajectory Analysis of Matched Percentiles (TAMP)?</u>

A Trajectory Analysis of Matched Percentiles (TAMP) is a powerful scale-independent technique that can be used to compare the change in test scores of groups of students (Braun, 1988). A TAMP permits a graphical assessment of whether one group of students exhibits "uniformly" or consistently slower progress relative to another group.

One of the major advantages of TAMP over more traditional analyses is the invariance to changes in scale. The comparison of test scores produces the same conclusions using TAMP under any order-preserving transformation of the test scores; only the ordinal relationship of the test scores matters. For this study, the scale invariance of TAMP permits a comparison of kindergarten test scores from the TOBE test with the CTBS scores (both converted to Expanded Scale Scores) obtained beginning in first grade. The scale invariance also allows analysis of the CTBS test scores without relying on the intended equal-interval properties of the Expanded Scale Scores (ESS) used in the growth curve analyses of Chapters III, IV, and V.

Another major advantage of TAMP is that it may be used to compare test score gains for groups of students even when they are not composed of exactly the same students. TAMP does not avoid the problems of trying to use cross-sectional data to measure growth. Growth is properly measured only with longitudinal data (see, for example, Rogosa, 1979; Rogosa, Brandt, and Zimowski, 1982; Rogosa and Willett, 1985; and Willett, 1989). This study uses longitudinal data, and individual student growth curves were calculated and compared in the analyses of earlier chapters. However, the two cohorts of late-exit students represent a cross-sectional sample. One group of late-exit students were in kindergarten in the first year of data collection; the other group of late-exit students were in third grade in the first year of data collection. Both groups were followed longitudinally and the growth of the individual students was evaluated. TAMP provides a mechanism for comparing growth across the two late-exit cohorts under the assumption that the two cohorts are similar. That assumption can be tested to some extent by comparing the third grade test scores from the two groups, one set obtained the first year of data collection and the other set obtained the fourth year of data collection.

How are the TAMP curves constructed for each group?

A TAMP curve, or "equipercentile equating function," for a group of students is a graph of ordered pairs of test scores. The pairs are constructed by associating each score in the distribution of scores from an earlier test with one score in the distribution of scores from a later test. The two scores that are associated have the same relative percentile. That is, the scores from the two tests are paired so that the scores in each pair have the same percentile rank within their respective distribution of test scores.

The two distributions of test scores need not be obtained from the same students, as long as all of the test scores are obtained from students who are representative of the same population. This permits TAMP curves to be constructed from cross-sectional data. For example, although the national norms for the TOBE and CTBS tests were constructed from cross-sectional samples, TAMP curves can be constructed because the group represented by the norms -- the U.S. grade-school population -- is well-defined.

For this project, the students were studied longitudinally. This permits the construction of TAMP curves from matched samples: the same set of students took both tests. Because the resulting curves are based on the same set of students, the analysis cannot be distorted by sampling fluctuations from the first test to the second test.

For matched samples, the construction of TAMP curves is particularly simple: the lowest score from the earlier test is matched with the lowest score from the later test, the next lowest score from the first test is matched with the next lowest score from the second test, and so on. This simplification arises because the two groups have the same number of students. To see this, suppose there are 200 students with scores in each of two grades. Then the 20th lowest score, for example, represents the same percentile point for both grades.

ERIC

All of the TAMP curves presented here are based on scores from matched students, except when comparing the K-3 and 3-6 late-exit cohorts. Although TAMP curves based on scores from all students who took each test have somewhat more stable estimates of percentile points, sampling fluctuations could produce distortions. When TAMP curves are based on all students who took either test, not just students who took both tests, the students included for each test are somewhat different samples of the limited-English-proficient (LEP) students in the district, program, or district-program. Any substantial difference between the all-students TAMP curve and the matched-students TAMP curve might be attributable to the differences in the two groups of students. Accordingly, the matched-students curves are preferable and are the ones presented here.

The TAMP curves presented here are often based on relatively few students. For this reason, the extreme percentile points are subject to considerable sampling error. The very lowest and highest scores are sometimes very far from the rest of the scores. To avoid having TAMP curve comparisons unduly influenced by the scores of a handful of students, only the middle ninety percent of the TAMP curves are presented here. That is, the highest five percent and the lowest five percent of the scores in each grade are not included in the figure. The middle part of the TAMP curves remains substantially more stable than either end, even after trimming the highest and lowest five percent.

<u>What do the TAMP figures look like</u>?

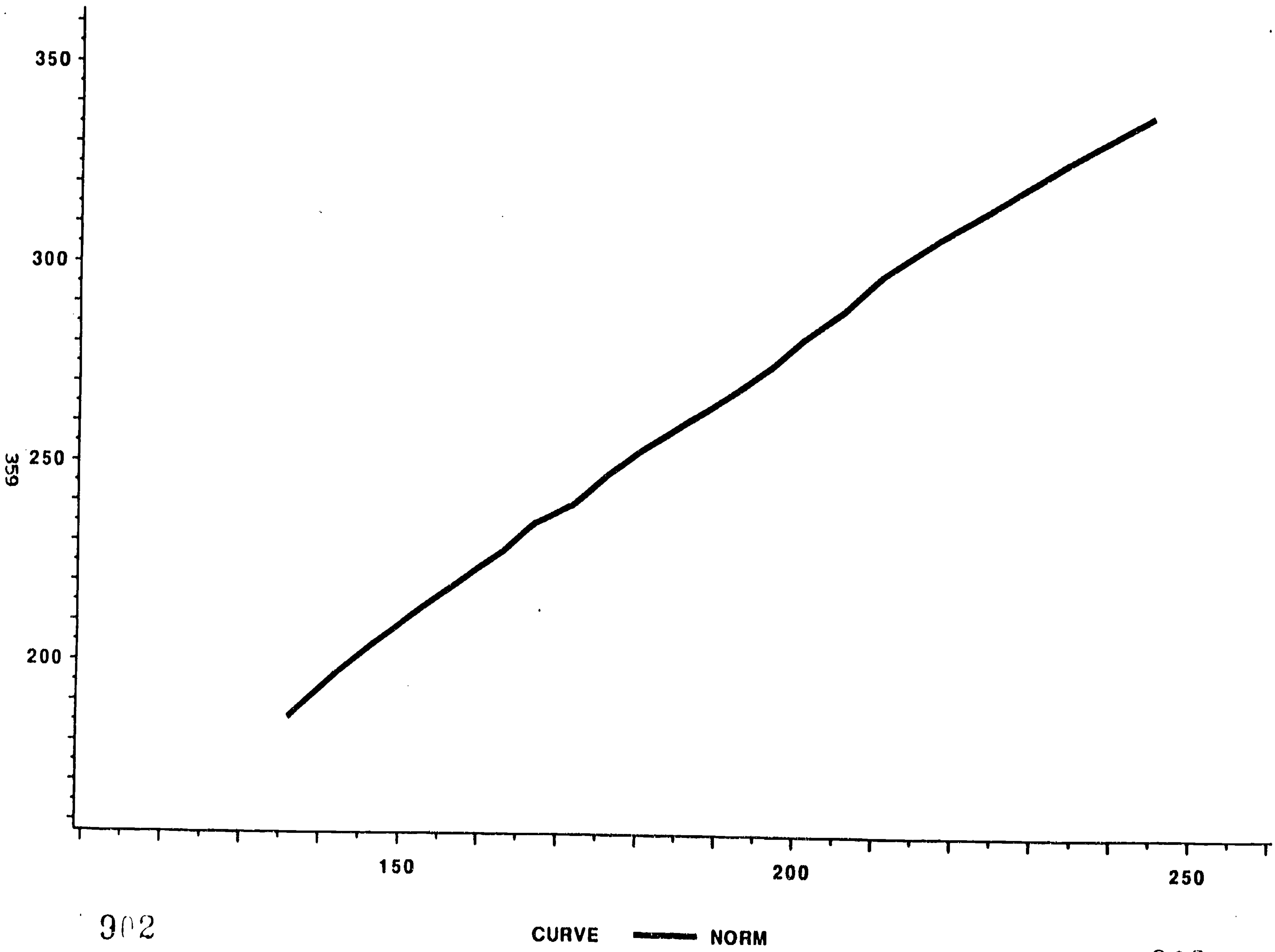
Each TAMP figure consists of one or more TAMP curves graphed on the same axes. The later test score is plotted using the vertical axis ("y-axis"), and the earlier test score is plotted using the horizontal axis ("x-axis"). In keeping with standard terminology, the label at the top of each figure indicates "later test score" versus "earlier test score." A second title line indicates which TAMP curves are included in the figure. The legend at the bottom shows the type of line used for each curve.

How is a TAMP curve interpreted?

Perhaps the simplest way to express the interpretation of a TAMP curve is that it represents the growth students would need to exhibit in order to maintain their relative standing in the group. Consider the TAMP curve in Figure 16, which relates the norms for the CTBS mathematics subtest as measured in English in spring of first grade with the norms for the TOBE mathematics subtest administered in Spanish in fall of kindergarten. (For a discussion of the national norms, see Chapter II.) Notice that it is possible to compare the test scores from two different tests using TAMP, even though the two tests may be on different scales.

The TAMP curve itself is nearly a straight line. This is not a requirement of TAMP, but rather a consequence of the fact that both the TOBE test scale and the CTBS test scale are designed to be approximately equal-interval. Each point on the curve represents a matched percentile. For example, the 50th percentile for the fall kindergarten TOBE mathematics test is 185 (Equivalent Scale Score, or ESS) and the 50th percentile for the spring first grade CTBS mathematics test is 260 (ESS). Thus a point is plotted at 185 on the horizontal axis and 260 on the vertical axis. In the TAMP figure, the horizontal axis always represents the earlier test and the vertical axis always represents the later test. Again, the fact that the 50th percentile is near the middle of the TAMP curve is a consequence of the test designs rather than a requirement of TAMP. As each of the percentiles in the norming information from each test is matched, another point is plotted. The points are then connected with a solid line. The individual percentile points are not distinguished in the figure.

English Math Spring 1 vs. Spanish Math Fall K



ERIC

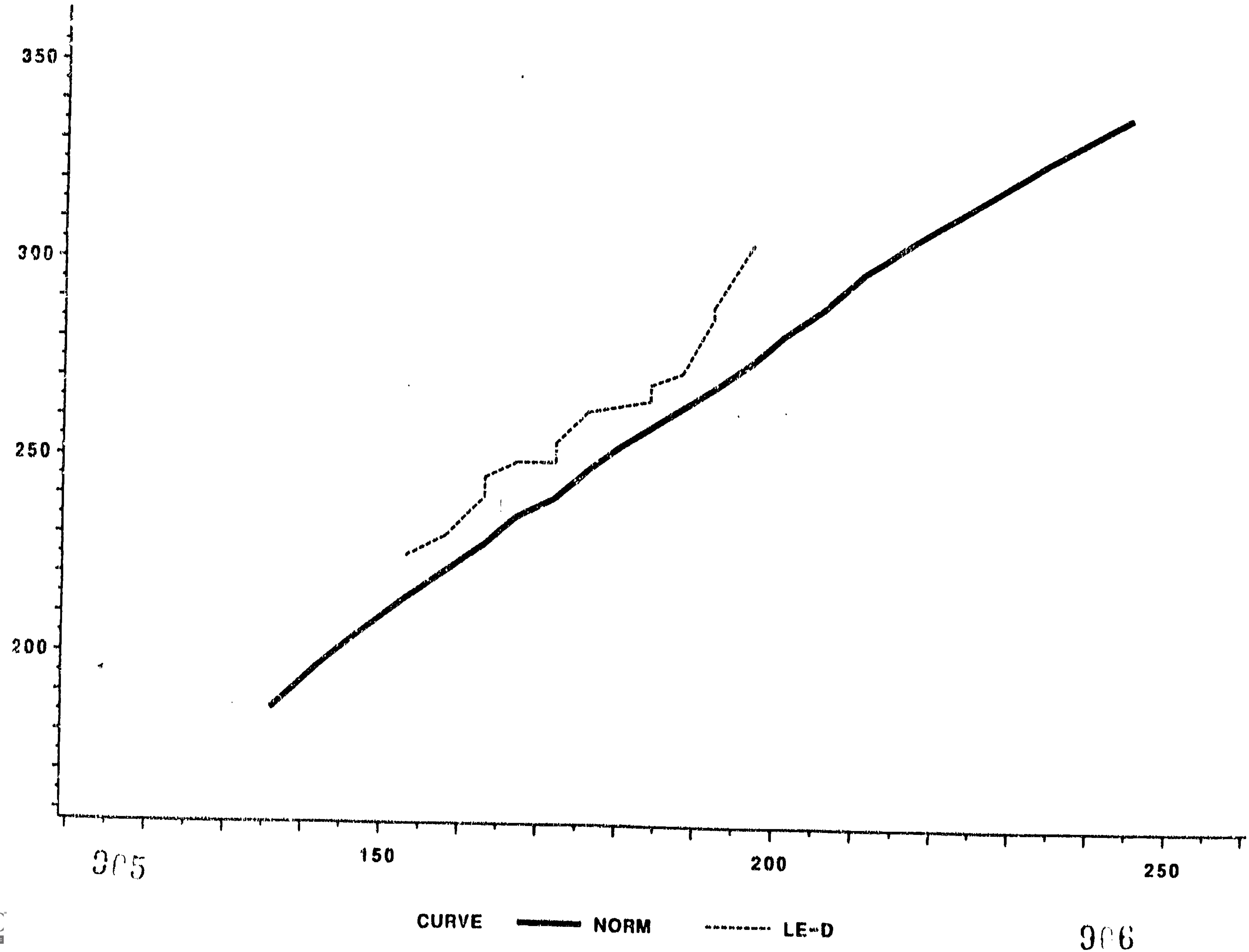
Thus each point on the TAMP curve shows (on the vertical axis) the test score students need to attain in order to score at the same percentile in spring of first grade as they did in the fall of kindergarten. A student scoring 185 on the fall kindergarten TOBE mathematics test is at the 50th percentile of the national norms for the TOBE. To score at the 50th percentile on the CTBS mathematics test in spring of first grade, a student needs to score 260 on the CTBS test.

How are two TAMP curves compared?

Two TAMP curves can be compared most easily on a graph. If the TAMP curve for group 1 lies entirely below the TAMP curve for group 2, then group 1 exhibits "uniformly retarded progress" compared with group 2. For an example of two groups where one displays uniformly retarded progress relative to the other, see Figure 17. The dotted line shows the TAMP curve for late-exit district D, while the heavy solid line is the same norm curve shown in Figure 16. The group represented by the national norms has uniformly retarded progress relative to the students in late-exit district D. To see why this is so, consider students scoring 170 on the fall kindergarten TOBE mathematics test administered in Spanish. To maintain their relative position among the students in late-exit district D, those students would need to score about 248 on the CTBS mathematics test in spring of first grade. To maintain their relative position among students in the national norm samples, students scoring 170 on the fall kindergarten TOBE mathematics test would only need to score about 238 on the spring first grade CTBS mathematics test. The same relationship between the two curves is obtained at every starting score represented in late-exit district D. Thus the students in late-exit district D exhibit greater growth for every starting score included in the figure.

Figure 17

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)



361

ERIC

It is tempting to interpret the range of the TAMP curve for district D as indicating the distribution of test scores in the district, with the middle of the TAMP curve representing the middle of the distribution of test scores. However, this is not a valid interpretation of the TAMP curve. While the range of the TAMP curve does show the range of the test scores after removing the lowest and highest five percent of the scores, the middle of the distribution of test scores may lie virtually anywhere along the curve. There may be many students near the low end of the range and only a few scattered along the rest of the curve. Or there may be many students near the high end of the curve and relatively few in the low and middle part of the curve. Or the test scores may be concentrated at a few points along the curve. In short, no inference may be drawn about the distribution of test scores along the TAMP curve.

One striking aspect of the TAMP curve for late-exit district D shown in Figure 17 is that it is not very smooth. TAMP curves are always monotonically nondecreasing. That is, as you move to the right along the curve it never goes down. A TAMP curve can, however, have horizontal or vertical segments. These are partly a consequence of the relatively small number of students used to construct the TAMP and partly a consequence of the coarseness of the test score distributions. Only 21 students in late-exit district D had pretest (fall kindergarten test) scores as well as first grade scores (see Chapter II). Because five percent of the scores were removed from each end of the distribution, the lowest two scores and the highest two scores were not included in the TAMP curve. Two scores were trimmed from each end because omitting only one score would have removed less than five percent of the scores; enough points were removed to represent at least five percent of the scores. Thus the TAMP curve is based on only 17 pairs of scores connected by straight line segments.

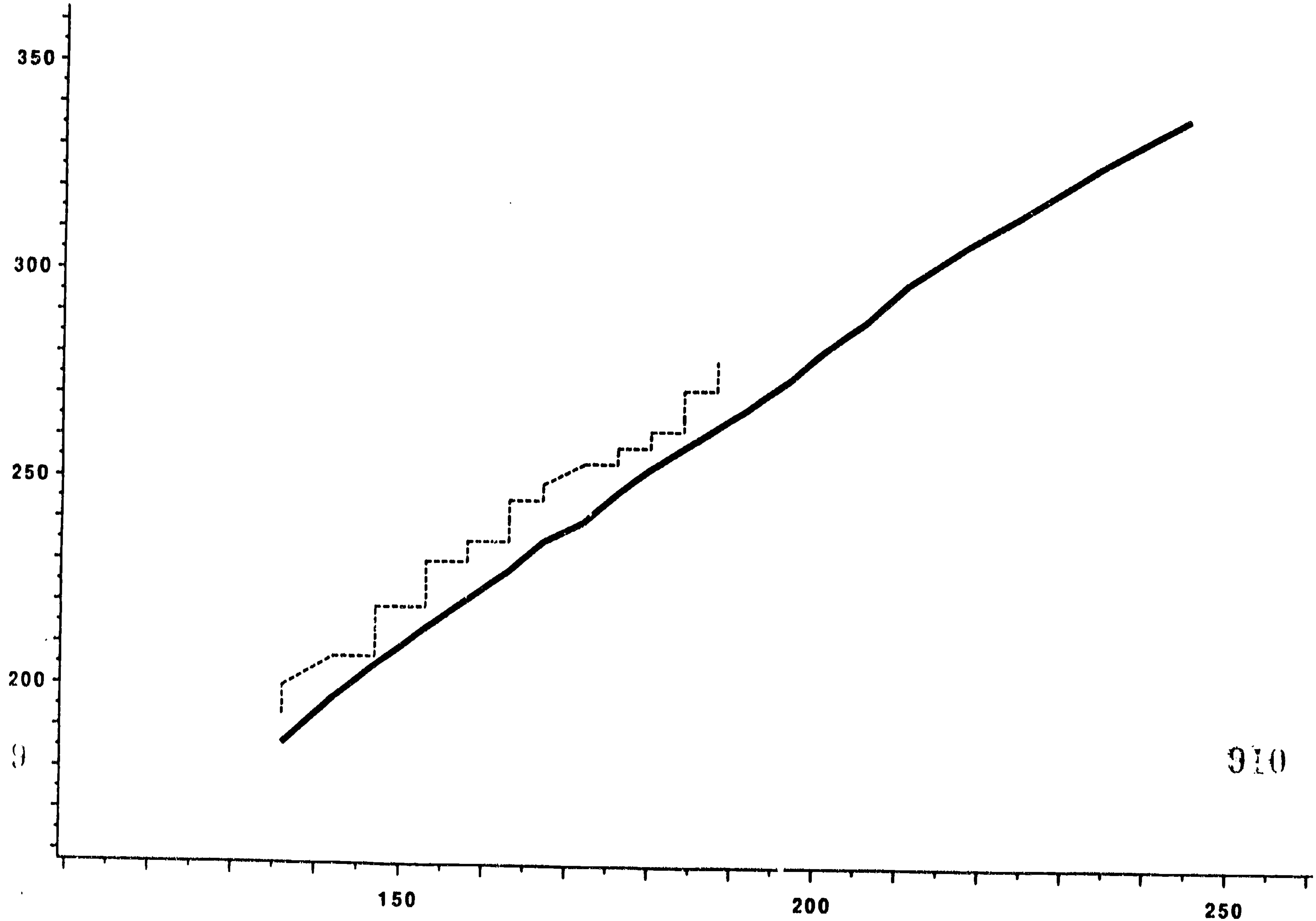
Even if many more scores had been used, however, the TAMP curve would still have been relatively jagged. This is because the tests contain relatively few items. The TOBE mathematics test, for example, contains only 26 items. Getting just one more item correct increases the scale score by several points. The CTBS test scores are similarly coarse.

Because only relatively few scores are possible on each test, the TAMP curve tends to take on a stairstep appearance. This is even more apparent in Figure 18, which shows the corresponding TAMP curve for late-exit district E together with the national norm TAMP curve. The students in this district also exhibit greater growth than the national norming samples (the TAMP curve for district E lies entirely above the national norm TAMP curve).

Both late-exit district D and late-exit district E show greater growth than the national norming samples. Figure 19 compares the two districts with each other. The two TAMP curves cross each other in several places. This means that neither district exhibits uniformly retarded progress relative to the other. Notice also that the two curves exhibit only partial horizontal overlap. The horizontal axis gives the score on the earlier test. The amount of horizontal overlap indicates the extent to which the two districts have similar ranges of test scores on the earlier test. Each curve represents the range of the middle ninety percent of the students in that district. Thus the overlap indicates the extent to which the range of test scores for the middle ninety percent of the students in district D is similar to the range of test scores for the middle ninety percent of the students in district E. As noted above, the TAMP curves do not indicate how the scores are distributed through that range.

Figure 18

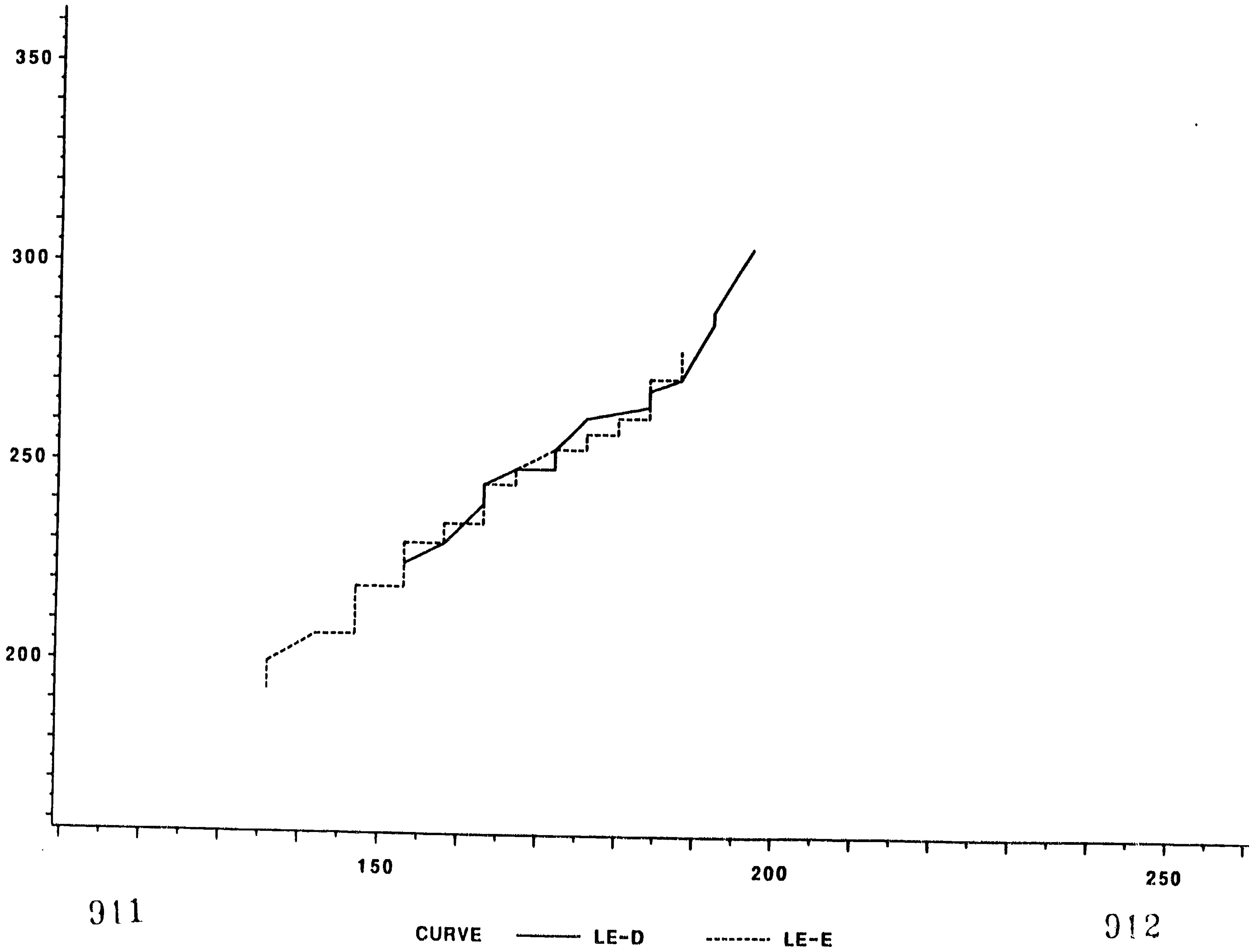
English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)



910

Figure 19

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)



365

911

912

ERIC

Two TAMP curves can only be compared over the part of the horizontal axis where they overlap because other parts of the curves correspond to students in one group with no counterparts in the other group. The district E students scoring below about 155 on the fall kindergarten TOBE mathematics test have no counterparts in district D. Since there are no students with comparable scores shown for district D, it is impossible to say how students scoring below 155 would have fared in district D. Similarly, the district D students scoring above about 187 on the pretest have no counterparts in district E. Again it is not possible to say how such relatively high-scoring students would have fared in district E. This reflects a strength of TAMP: group comparisons are necessarily restricted to students with similar scores on the earlier test.

The importance of the finding that neither district D nor district E exhibits uniformly retarded progress relative to the other depends on whether a difference was expected. If one district placed much more emphasis on mathematics, the students in that district would be expected to show greater progress, and the absence of demonstrably higher growth would be a disappointment. However, because there is no particular reason to expect a difference between these two districts, the TAMP finding is unremarkable.

Are the TAMP curves adjusted for any differences between groups of students?

The TAMP curves are not adjusted in any way for group differences on any variable other than the score on the earlier test. Therefore, differences in socioeconomic status (SES) of the families in one district might account for any differences in growth. While it is technically possible to produce separate TAMP figures according to some measure of SES or other covariates thought to predict school achievement, no covariate-adjusted TAMPs were performed for this study due to resource limitations. The number of separate TAMP figures for comparing the districts, programs, and district-programs at the various grades is already too large for all of them to be presented here.

ERIC

Because the TAMP curves are not adjusted for any background variables, the comparisons of the districts and programs using TAMP must be made in the context of the many differences among the districts. The TAMP figures cannot answer whether any growth differences between two districts (an IS/EE district and an LE district, for example) are attributable to differences in student, school, or district characteristics or are attributable to the bilingual program. With so few districts, conclusions about the possible explanations for any district differences must be made on a case-by-case basis, not by using statistical analyses.

How are TAMP curves used to compare the K-3 and 3-6 late-exit cohorts?

Matched-students TAMP curves are not possible for comparing the early grades in the late-exit districts with the later grades. This is because the two late-exit cohorts, one covering grades K-3 and the other covering grades 3-6, are made up of different students. By assuming that both cohorts are representative of the district LEP student population, TAMP curves can be constructed showing growth across the two cohorts. A comparison of first grade to sixth grade could be constructed, for example. Because the resulting curves are only appropriate if the two late-exit cohorts are comparable, the comparisons in Chapter V of the two cohorts at third grade should be considered in connection with the TAMP figures.

What are some of the limitations on the interpretation of the TAMP figures?

Because the national norms are based on much larger samples than the curves based on data from this study, all but the single highest and lowest percentile points appear to be reliably estimated. Therefore, the TAMP curves for the norms represent very nearly the entire range of percentiles from 1 to 99.[3] For the purposes of the TAMP figures presented

[3]The TAMP curves for the national norms range from an NCE (Normal Curve Equivalent) of 2 through an NCE of 98.

ERIC

here, the national norm curves can be viewed as having essentially no sampling error. However, it should be pointed out that the TOBE norms are based on a 1977-78 sample, while the CTBS norms are based on a 1972-73 sample. (This relatively old version of the CTBS test was used at the request of the U.S. Department of Education to ensure comparability with other national studies.) These samples, therefore, are from different national populations, and neither reflects the national population in 1984 to 1988, when the test scores in this study were collected. (For more information on the TOBE and CTBS tests used, see Chapter II.) Only if there are no time trends in the distribution of test scores for this national population would the TAMP curves for these norms represent a completely accurate reference curve. As noted in Chapter II, there is some evidence that such time trends might exist (Linn, Graue, and Sanders, 1990).

Braun (1988) gives an approximate formula due to Lord (1982) for the standard error of TAMP curves. This formula can be used to provide some guidance about precisely how the TAMP curves are estimated. Naturally, the standard error varies considerably according to the number of students in the group used to construct the curve. Most of the curves presented in this chapter are based on one hundred or more students, but some are based on only a few dozen, or even fewer. The curves for late-exit district D are based on the fewest students and therefore have the highest standard errors.

The standard error of TAMP curves also depends on the underlying standard deviation of the test scores being compared. The standard deviations of the achievement tests used here vary considerably according to the subtest and the grade level. However, the TAMP figures are constructed so that one standard deviation represents approximately the same distance on each figure. Therefore the standard error of the TAMP curves can be expressed most easily in terms of distance on the graphs.

On the scale of these graphs, one standard error for the ends of a TAMP curve ranges from about one-half inch (for the smallest groups) to about two-tenths of an inch (for the largest groups), or even less. The standard error in the middle of the TAMP curves is about two-thirds as large as the standard error at the ends, so a standard error in the middle ranges from about three-tenths of an inch down to about one-tenth of an inch. Curves need to differ by about twice the standard error of the less precise curve to be considered appreciably different.

Organization of TAMP Results

The discussion of the TAMP figures that follow are divided into two sections. The first section is a discussion of the TAMP figures wherein the growth in achievement for the immersion strategy, early-exit and late-exit programs is compared to the growth of this norming population, as well as examining the growth in achievement for each of the immersion strategy and early-exit sites relative to this norming population. This first section is subdivided into four parts. Part 1 relates the academic growth rate of students in each program to that of this norming population from fall to spring of their kindergarten year. This addresses the question of whether the initial skills of target students in each program were comparable to one another and to those of this norming population, as well as the extent to which the students in the three programs had different achievement growth rates during the kindergarten year. Part 2 relates the growth in achievement of target students to this norming population from kindergarten to first grade, by using their spring kindergarten test scores and their spring first grade test scores, while Part 3 spans the students' kindergarten and first grade years by using their fall kindergarten scores and their spring first grade scores. Parts 2 and 3 address the issue of the extent to which the students in each program have different spring first grade scores, and therefore correspond to the K-1 analyses of Chapters III through V. Part 4 examines the relationship of target students' spring first grade to spring third grade test scores to those of this norming population. Part 4 expands the analyses presented in Chapters III through V by relating the academic growth of target students to that of this norming population; that is, the extent to which the target students are growing as fast as or faster than students in this norming population. This information will help us identify how successful each program is in helping students approximate the distribution of the norming population.

The second section provides a more detailed look at the growth rate of each of the late-exit program sites relative to the growth rate of this norming population. This is done in response to the unexpected finding in Chapter V wherein, contrary to predictions, the form of the late-exit students' growth curves in mathematics, language, and reading did not differ from the form of the growth curves for immersion strategy and early-exit students. This section examines the relationship of the three late-exit districts for each of the same comparisons done in Parts 1 to 4 of the first section. In addition, this section includes a comparison of academic growth spanning grades one to six for target students and for the norming population by merging the TAMP curves of the K-3 cohort with those of the 3-6 cohort. This expands the discussion provided in Chapter V regarding the growth of late-exit target students to include a comparison with national norms.

This second section is subdivided into six parts. Part 1 relates the academic growth of target students in each late-exit site to that of this norming population from fall to spring of their kindergarten year. This allows us to determine whether the initial skills of target students in each late-exit site were comparable to one another. This analysis also permits us to determine the extent to which the different implementations of the late-exit model are related to the growth in skills during the kindergarten year. Part 2 relates the academic growth of target students to growth for this norming population from kindergarten to first grade, by using their spring kindergarten test scores and their spring first grade test scores, while Part 3 spans the students' kindergarten and first grade years by using their fall kindergarten scores and their spring first grade scores. These sections address the issue of the extent to which each late-exit treatment has different spring first grade scores, and therefore correspond to the K-1 analyses of Chapter V. Part 4 examines the growth of target students' spring first grade to spring third grade test scores

ERIC

relative to that of this norming population. This section expands the analyses presented in Chapter V by relating the growth in achievement of target students to the growth of this norming population; that is, how successful each late-exit site is in helping move the distribution of scores for limited-English-proficient students who are below average to approximate that of this norming population. Part 5 includes a comparison of student academic growth spanning grades three through six, while part 6 includes a comparison of student academic growth spanning grades one to six for target students and for this norming population by creating TAMP curves using students from the K-3 cohort and students from the 3-6 cohort.

The following paragraphs comprise guidelines for reading and interpreting the figures included in this chapter:

If the TAMP curve for a given program is above the TAMP curve (greater than 1/2-inch) for this norming population, this implies that target students are growing at a faster rate relative to this norming population and are catching up as a group to this norming population. For example, the percentile rank of the second test score for target students is greater than the percentile rank of their initial test score.

If the TAMP curve for a given program is on the TAMP curve (less than 1/2-inch) for this norming population, this implies that the target students are maintaining their position relative to that of their counterparts in this norming population. For example, the second percentile rank for target students is equal to their initial percentile rank.

If the TAMP curve for a given program is below the TAMP curve (greater than 1/2-inch) for this norming population, this implies that target students are losing ground relative to this norming population. For example, the second percentile rank for target students is lower than their initial percentile rank. This suggests that students are not learning as fast as this norming population.

The Growth of Immersion Strategy, Early-Exit, and Late-Exit Students Relative to the Growth of This Norming Population

1. Fall Kindergarten to Spring Kindergarten:

A.				
	Grade Span:	**Kindergarten**	**to**	**Kindergarten**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**Spanish**
	Content:	**Mathematics**	**to**	**Mathematics**

Figures 20, 21, and 22 suggest that target students in all three programs had lower mathematics skills as assessed in Spanish upon entry into kindergarten than this norming population, as the spread of scores for each program tends to fall at the lower end of the norming curve. As a group, immersion strategy students appear to have lost ground; they seem to have been learning mathematics more slowly than this norming population when tested in Spanish. Early-exit and late-exit students seemed to grow at the same rate as this norming population.

Minor variation is noted when the TAMP curves are examined by site within the immersion strategy and early-exit programs. Whereas students in immersion strategy sites IS-B and IS-C seemed to have grown at a slightly slower rate than this norming population and site IS-H at about the same rate, students in site IS-A grew at a noticeably slower rate than this norming population, especially those students with the lowest initial entry-level skills (see Figures 23 to 26). All early-exit students seem to have grown at or almost at the same rate as this norming population (see Figures 27 to 30). Sites EE-A and EE-H appear to have done a better job of helping their students, as those with the lowest initial mathematics skills grew as fast as this norming population. In site EE-C students with the lowest initial skills seemed to grow more slowly than this norming population, while those at the high end grew faster; the distribution of scores is stretching out.

Contrary to expectations, there is some indication that early-exit students performed slightly better than did immersion strategy students in relationship to this norming population when tested in Spanish. It appears that early-exit students grew at about the same rate as this norming population. Given that early-exit students were not expected to receive much instruction in Spanish and what was provided was limited to initial literacy skills in Spanish, it had been predicted that these students would perform as well as immersion strategy students, and that both of these groups when re-tested in Spanish would post only minimal gains. However, early-exit students with higher entry-level mathematics skills appeared to be learning at the same rate or even slightly faster than this norming population, whereas their immersion strategy counterparts seemed to be growing slightly slower than this norming population. Also, early-exit students with the lowest entry-level skills appeared to be learning at a slightly faster rate than comparable immersion strategy students relative to this norming population.

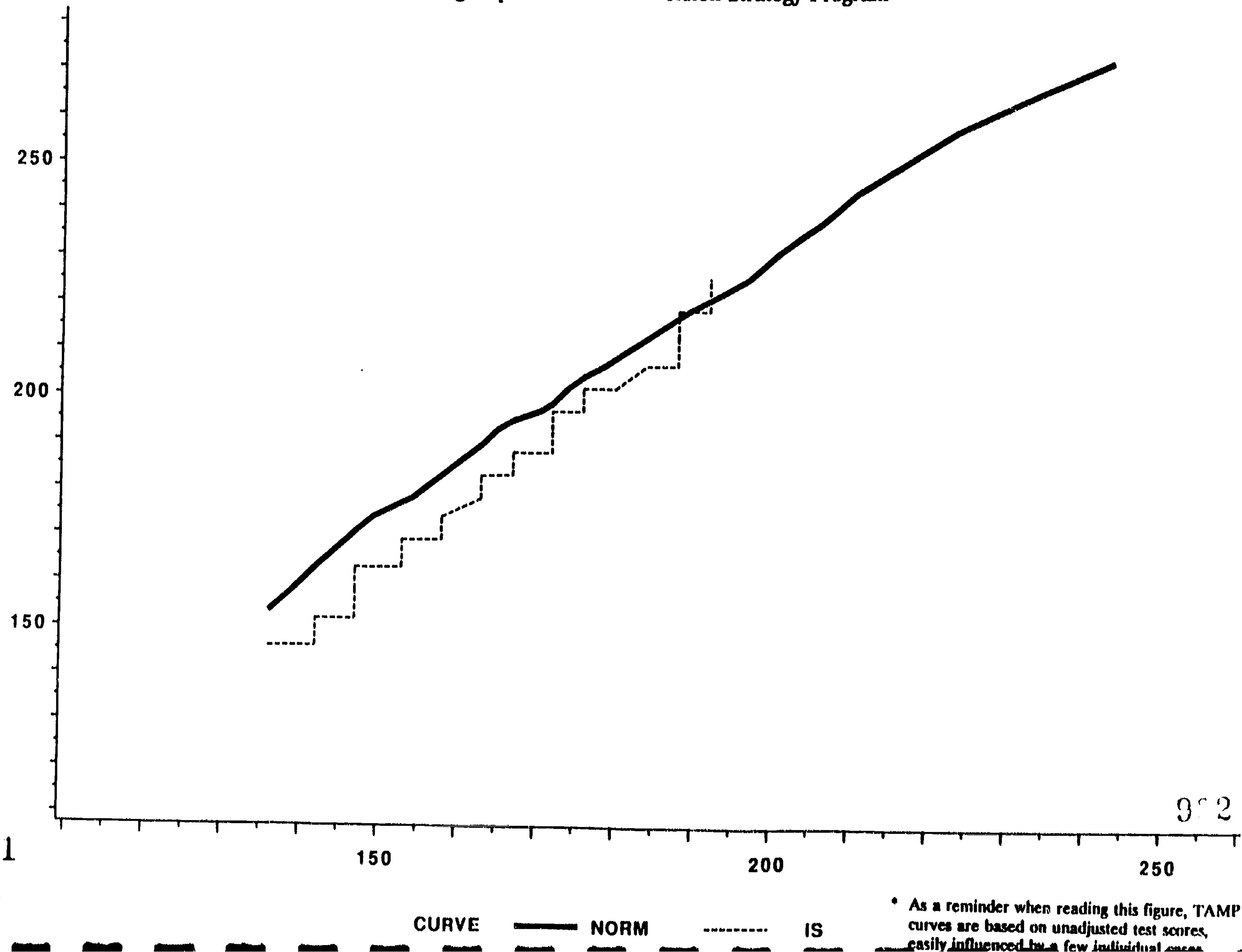
In sum, immersion strategy students appear to have lost ground relative to this norming population, whereas their late-exit and early-exit counterparts were doing as well as or slightly better than this norming population. The slightly better performance of the early-exit students may reflect the finding that some early-exit programs did in fact provide more instruction in Spanish than had been projected.

ERIC

Figure 20

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

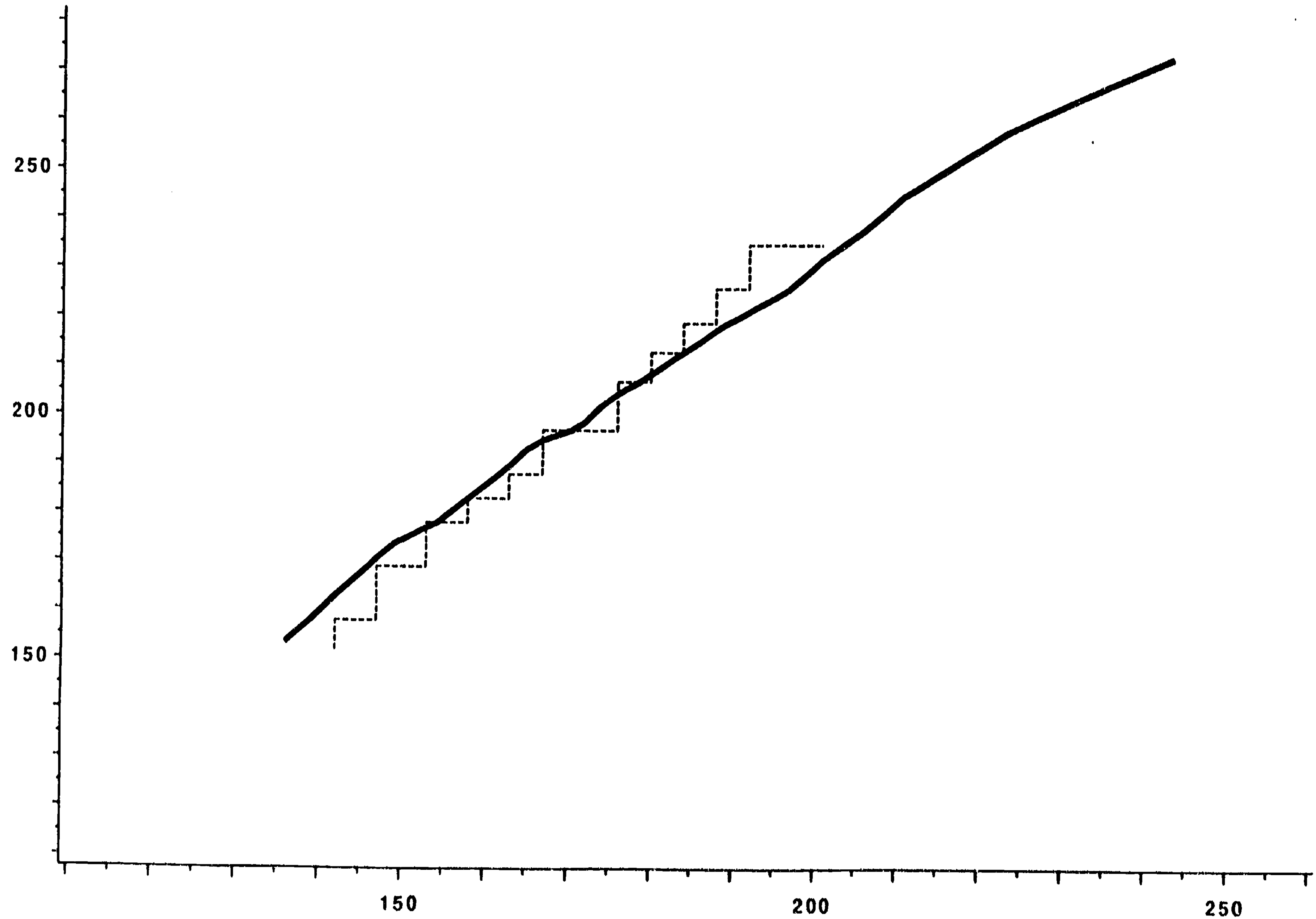


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 21

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



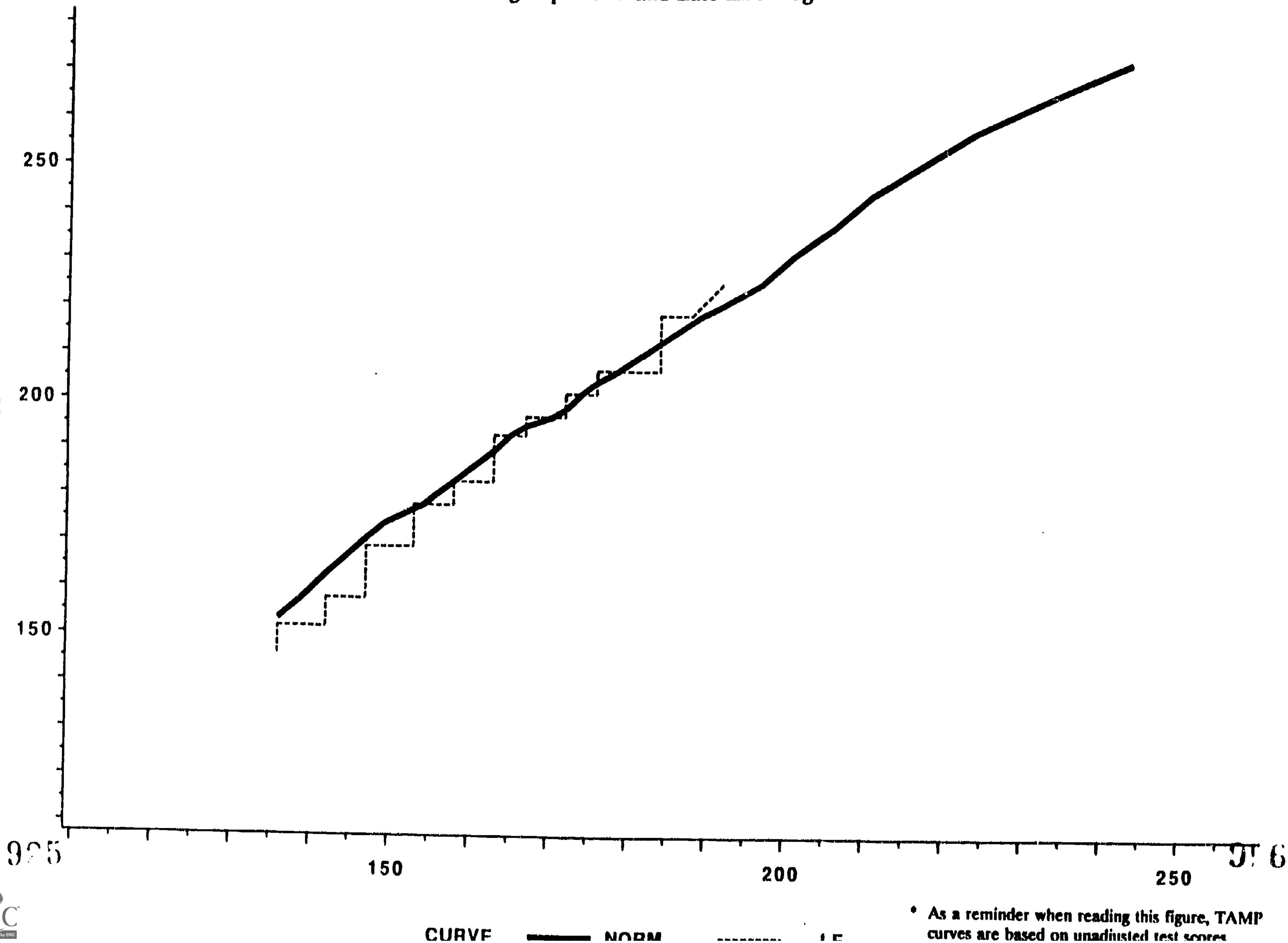
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 22

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



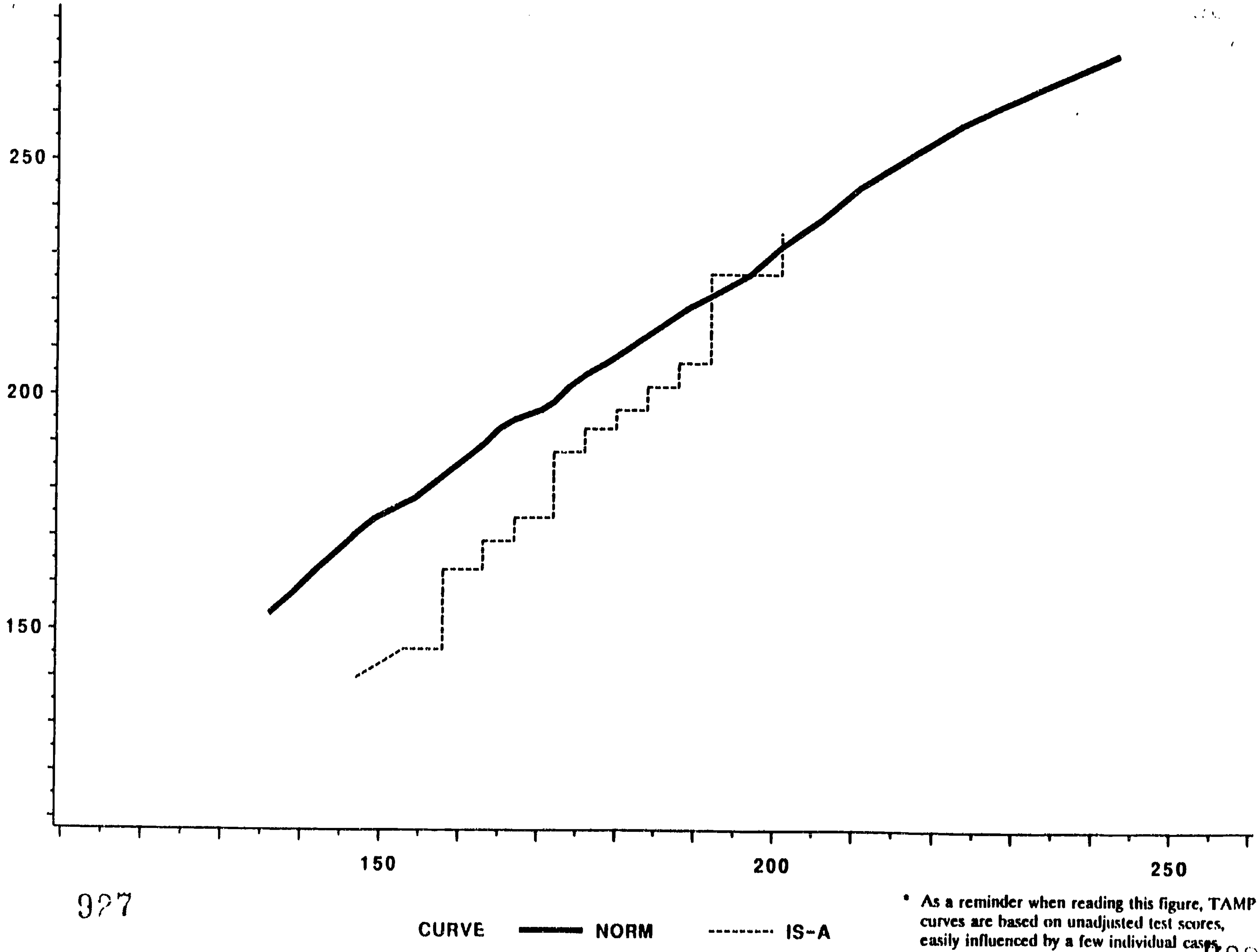
376

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores

Figure 23

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



377

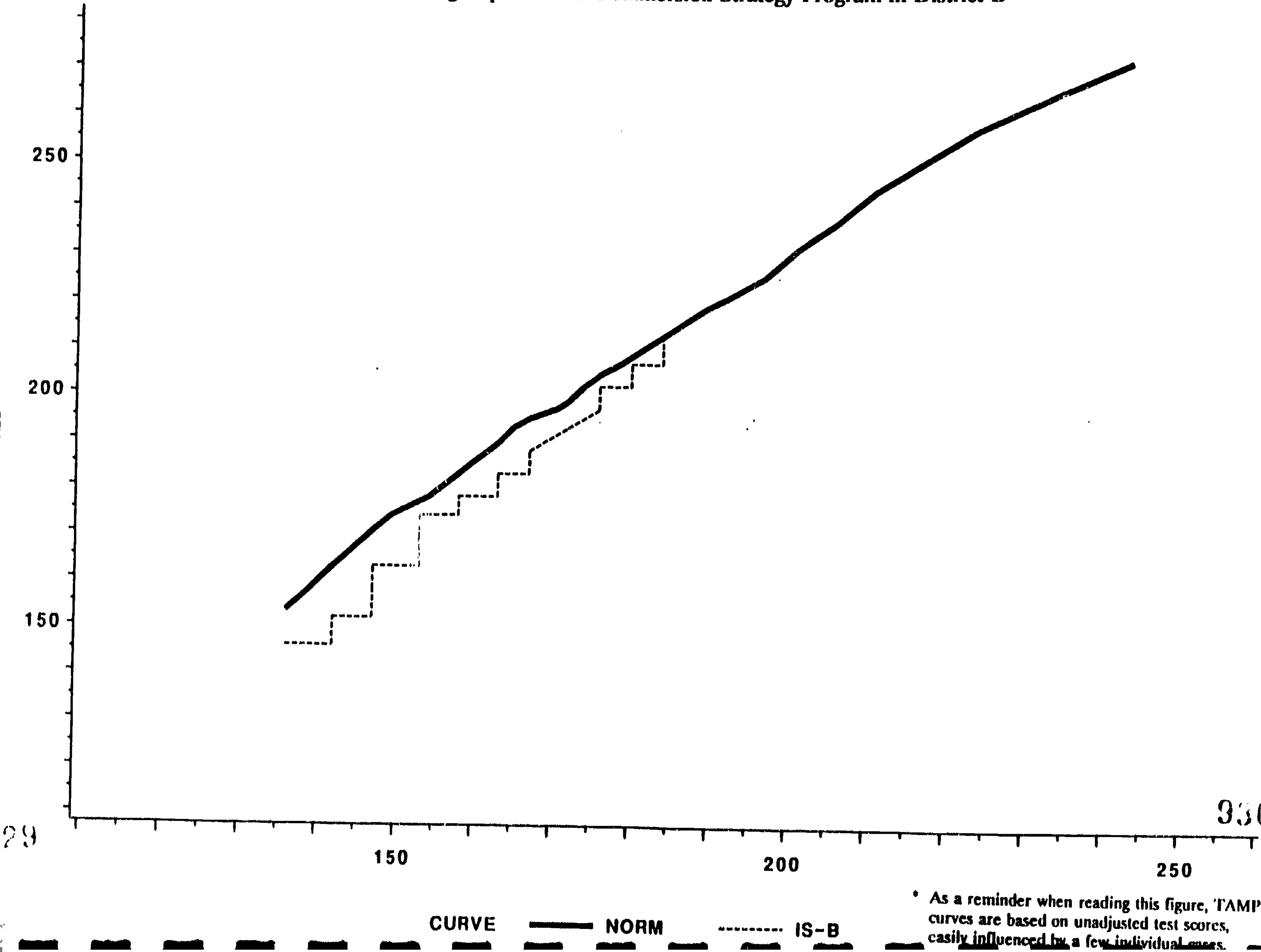
927

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 24

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



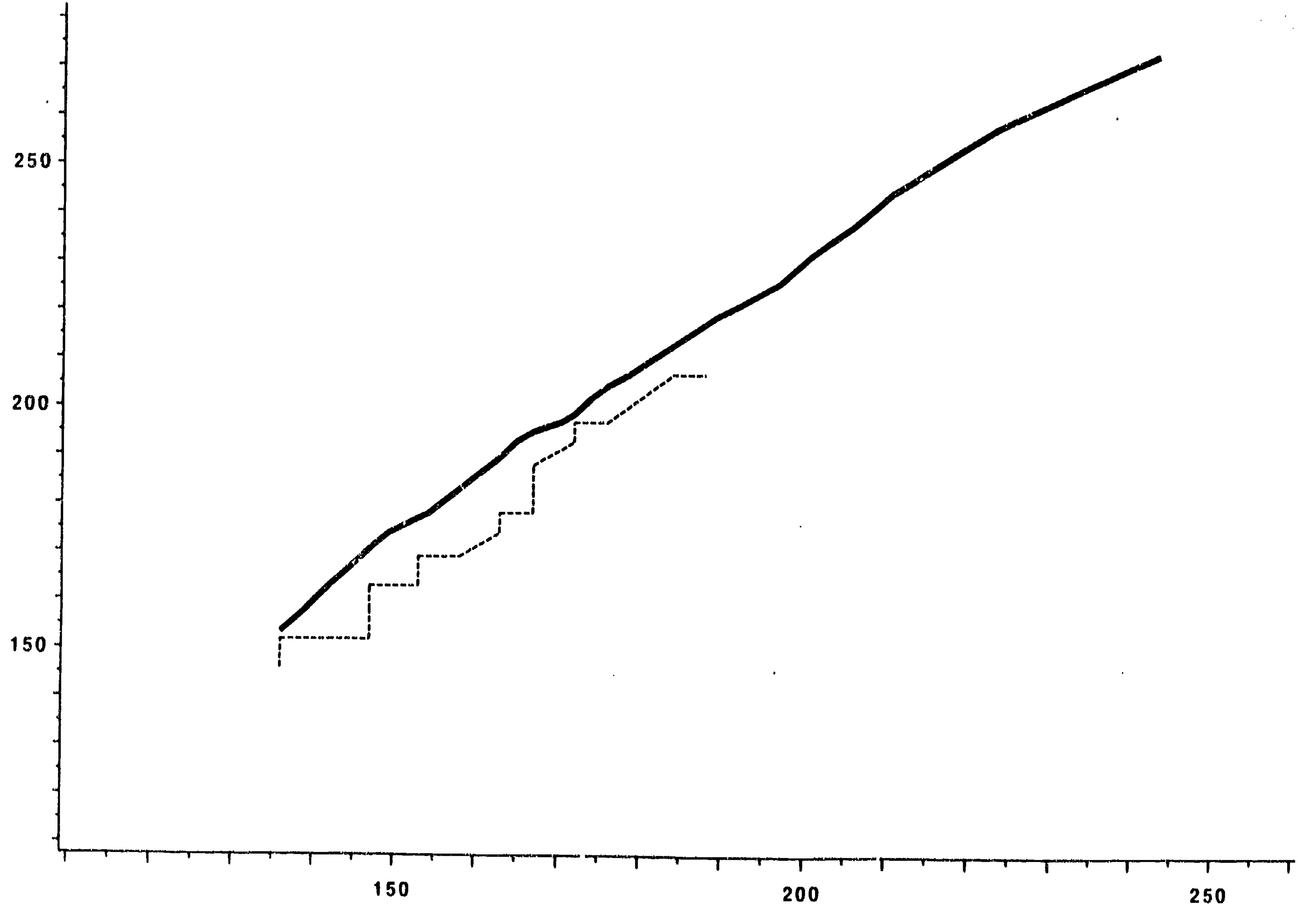
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, casily influenced by a few individual cases.

ERIC

Figure 25

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C

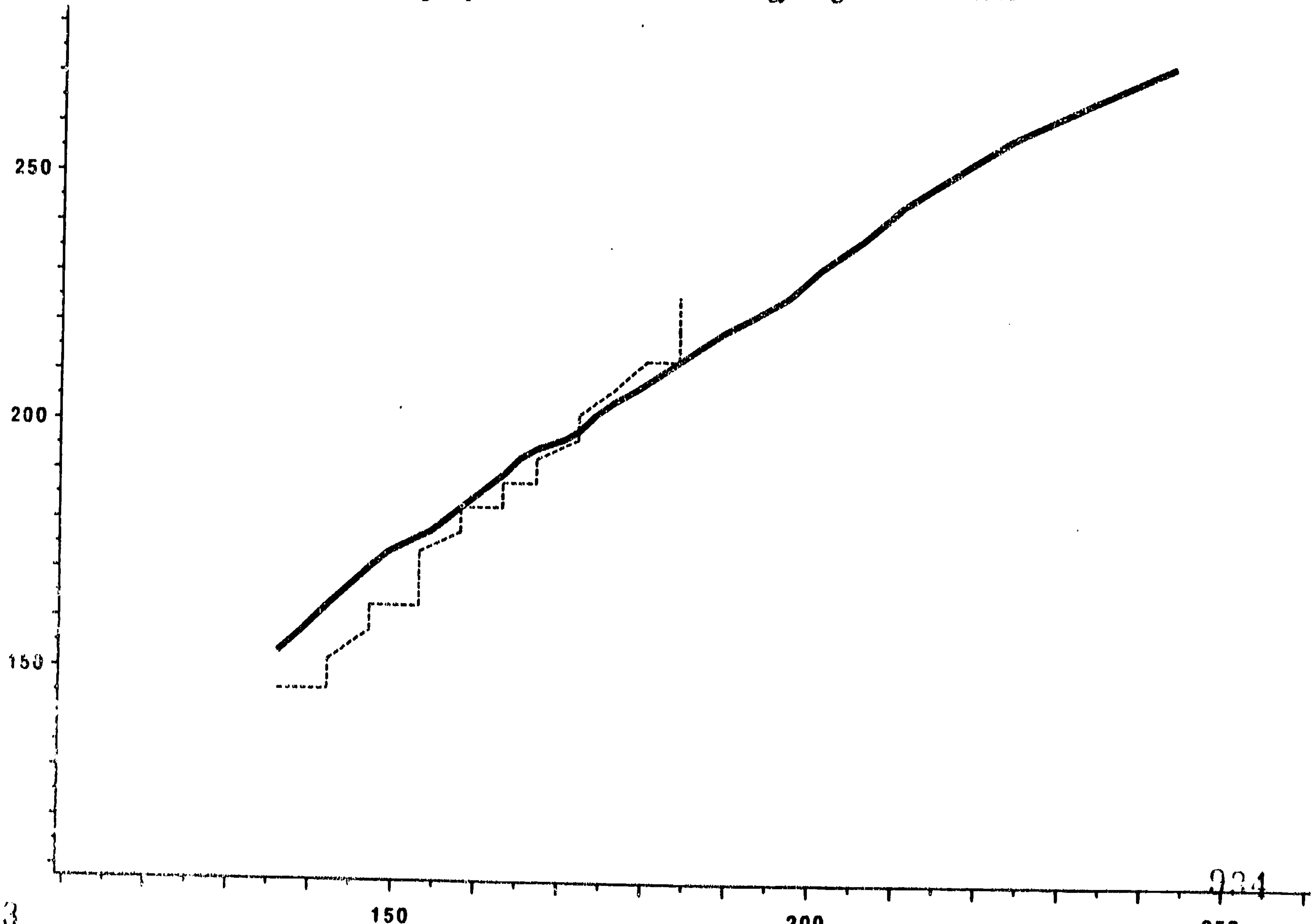


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 26

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H

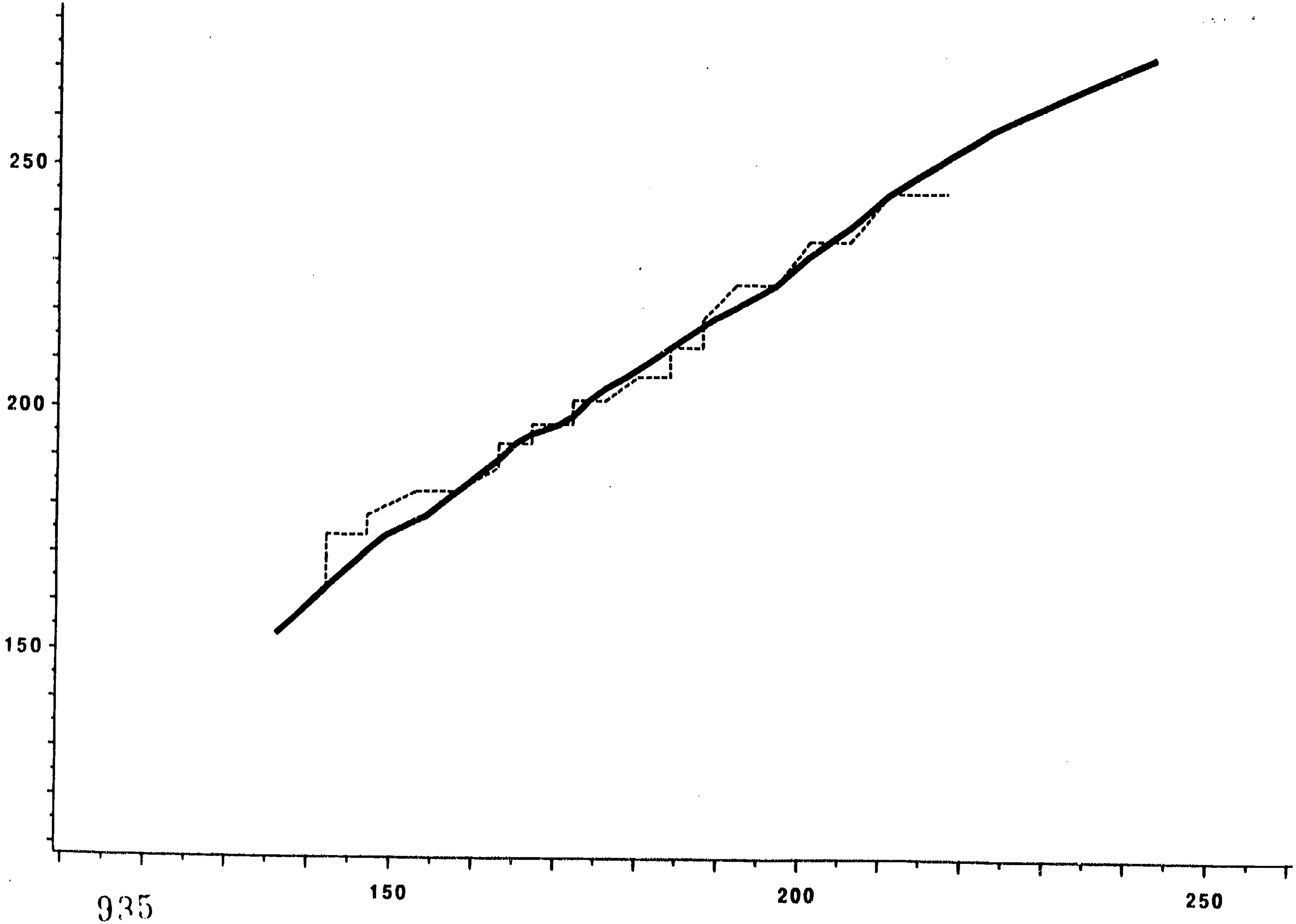


* As a reminder when reading this figure, TAMP

Figure 27

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A

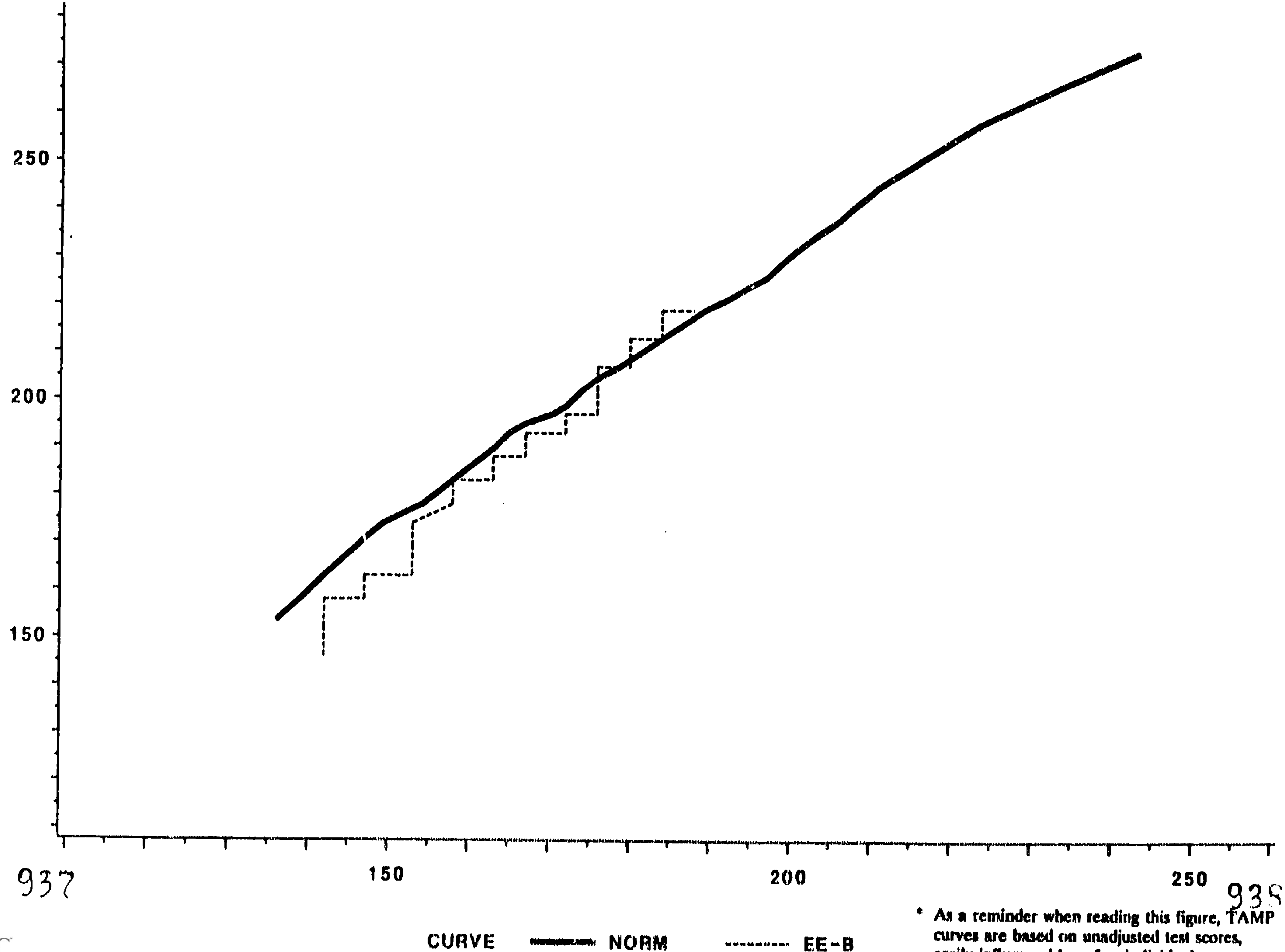


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 28

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



382

937

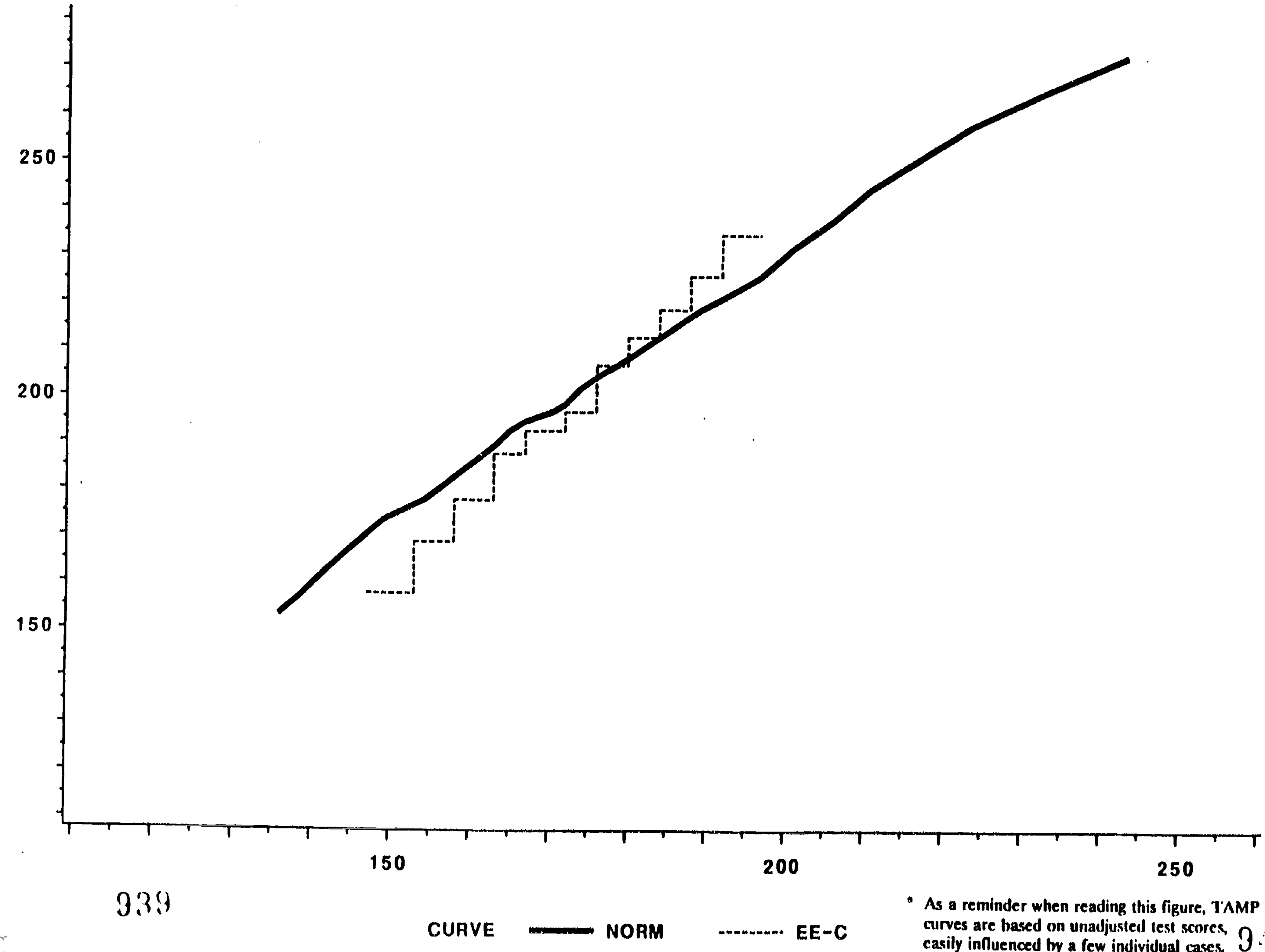
938

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 29

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



383

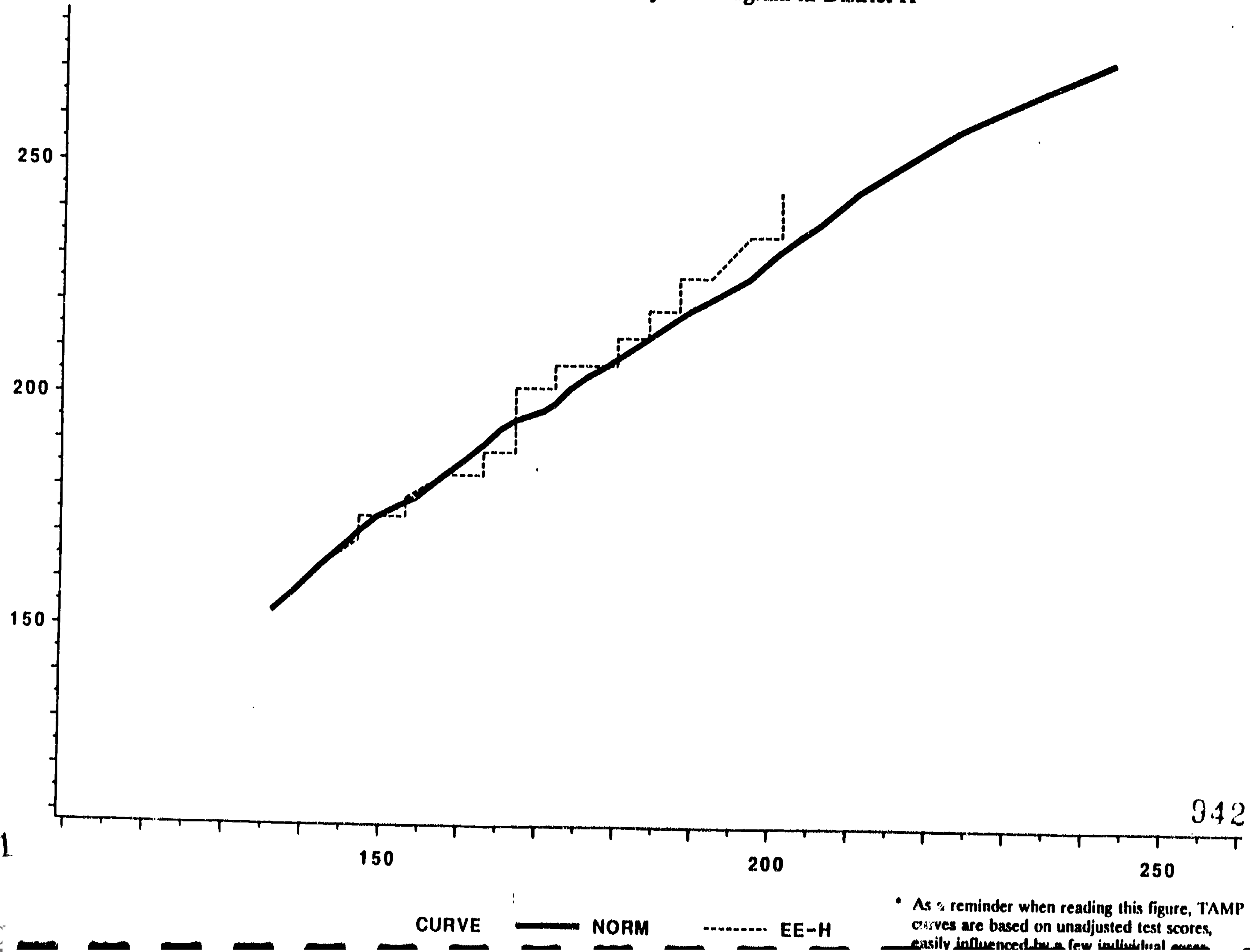
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 30

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual scores.

B.	**Grade Span:**	**Kindergarten**	**to**	**Kindergarten**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

Figures 31, 32, and 33 compare the mathematics test scores in fall kindergarten with the mathematics test scores in spring kindergarten. Both the fall and spring test scores are from the TOBE, but the fall TOBE was administered in Spanish and the spring TOBE was administered in English. For the TOBE, the test items are the same regardless of the language of administration; only the language used for giving the instructions varies. Because the same test was administered in fall and spring, comparisons across languages of administration are reasonable. It should be kept in mind that only the ordinal relationship of the test scores matters for TAMP.

The national norm TAMP curves presented in this chapter for the TOBE are the same for both languages of administration. The national norms reflect administration of the TOBE in English primarily to monolingual English-speaking students. It is difficult to relate these national norms to expectations for the limited-English-proficient students in this study. The students in the national norming sample were given test instructions in English in both fall and spring; the target students were given test instructions in Spanish in the fall but were given the test twice in the spring, once with instructions in Spanish and once with instructions in English. Figures 31, 32, and 33 reflect the spring administration of the mathematics subtest using English instructions. Presumably the students in the national norming sample have an advantage because they had both tests administered in English, their primary language, and because their instruction was all in English. The magnitude of this advantage is unknown.

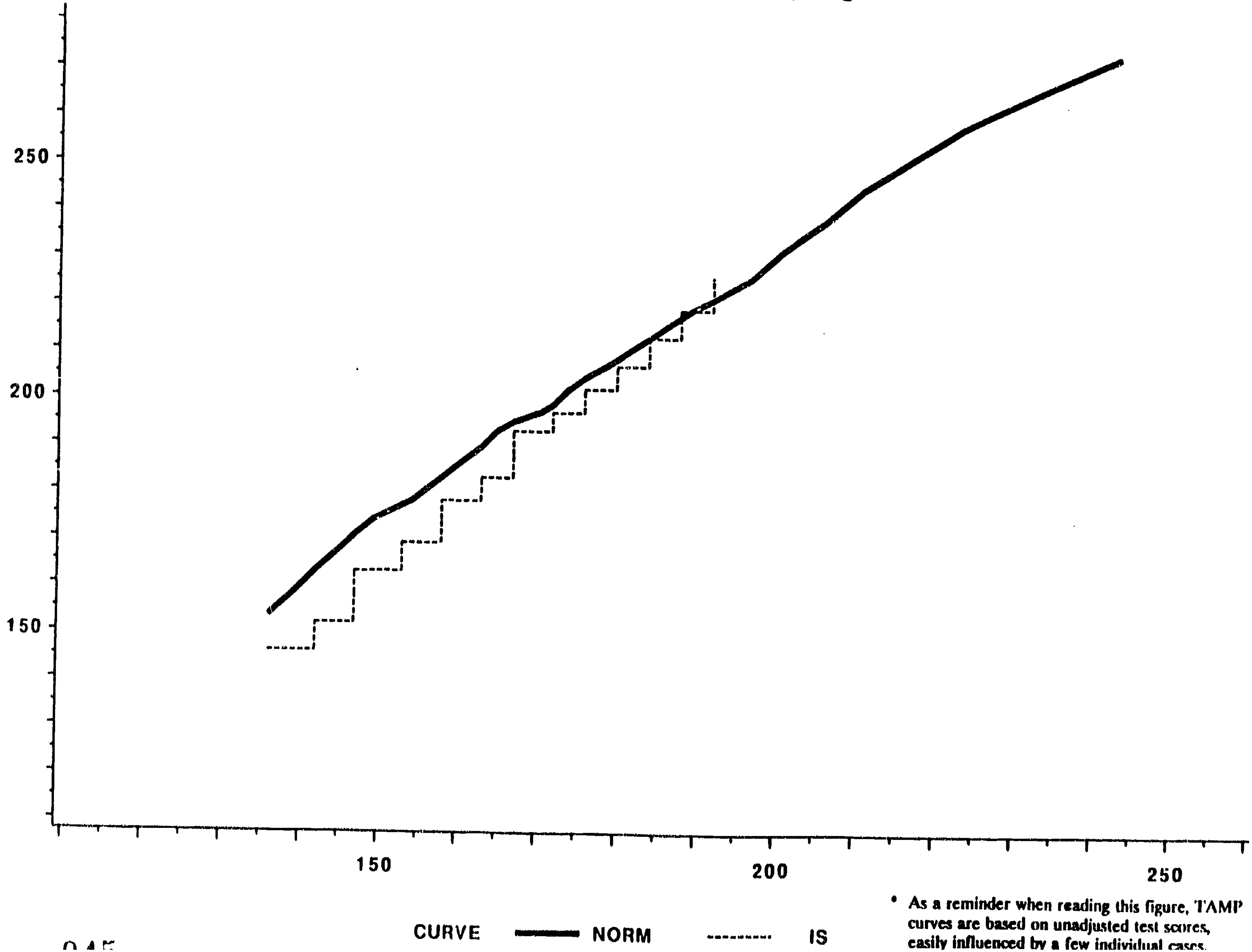
Note that the distributions of scores in Figures 31, 32, and 33 for each of the three programs tend to be shifted to the lower left of the norming curve. This indicates that in general the majority of the target students in all three programs began kindergarten with mathematics skills comparable to one another but lower than those of this norming population. The TAMP curves suggest that the growth in mathematics skills is similar in early-exit programs in relation to this norming population. That is, target students in each program appear to have kept pace with the growth of this norming population. However, immersion strategy and late-exit students may have been losing some ground, especially at the lower end of the distribution.

Some within-program variation was noted in both the immersion strategy program sites and the early-exit program sites. Figures 34 through 37 appear to show that while students in sites IS-B and IS-H grew at the same rate as this norming population, students in sites IS-A and IS-C grew more slowly. Figures 38 through 41 suggest that while students in sites EE-A and EE-H were comparable in their rate of growth (growing slightly faster than this norming population), site EE-B students seem to be about even with these national norms, and site EE-C students were losing ground in mathematics as measured in English relative to this norming population.

Figure 31

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

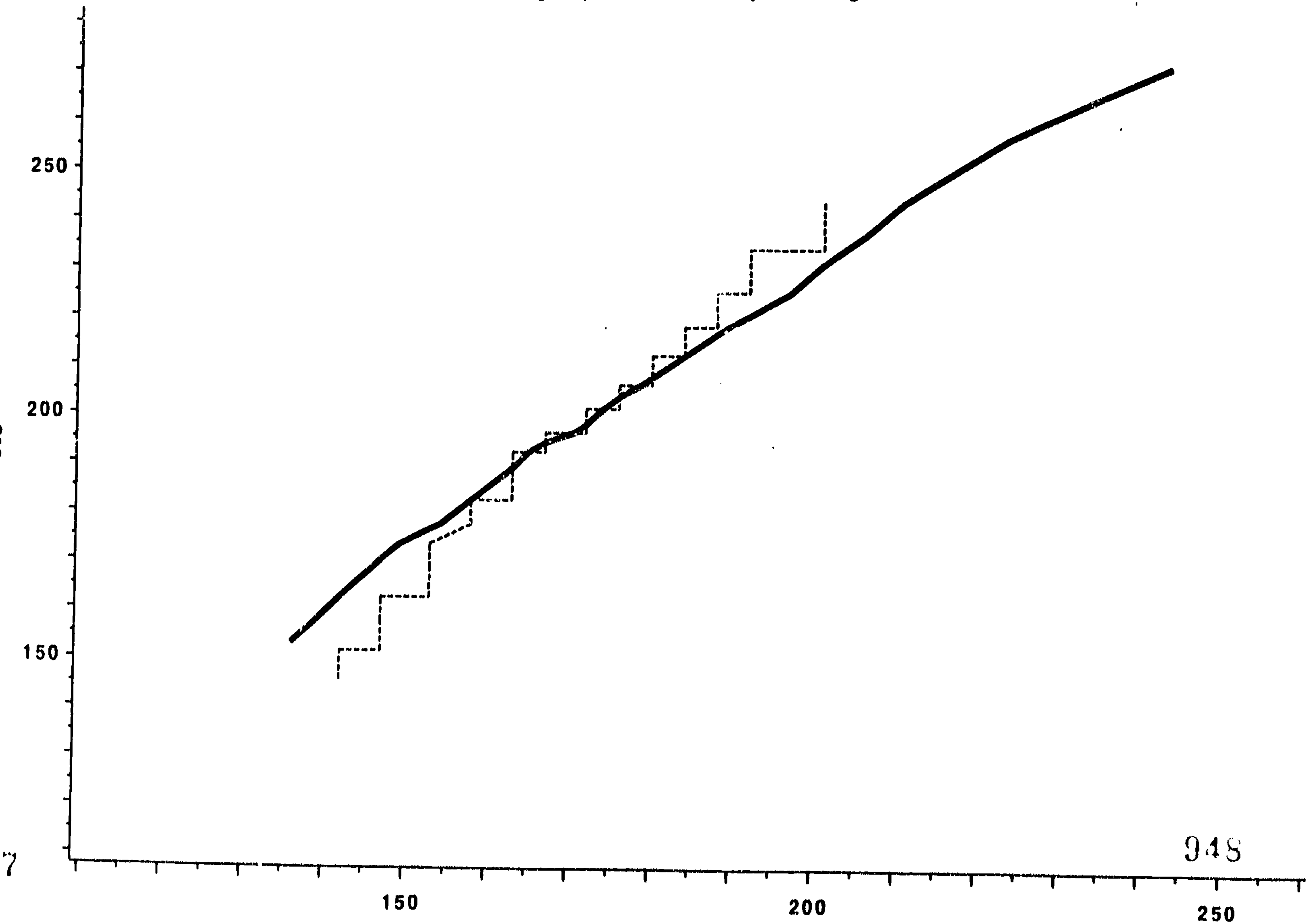


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 32

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



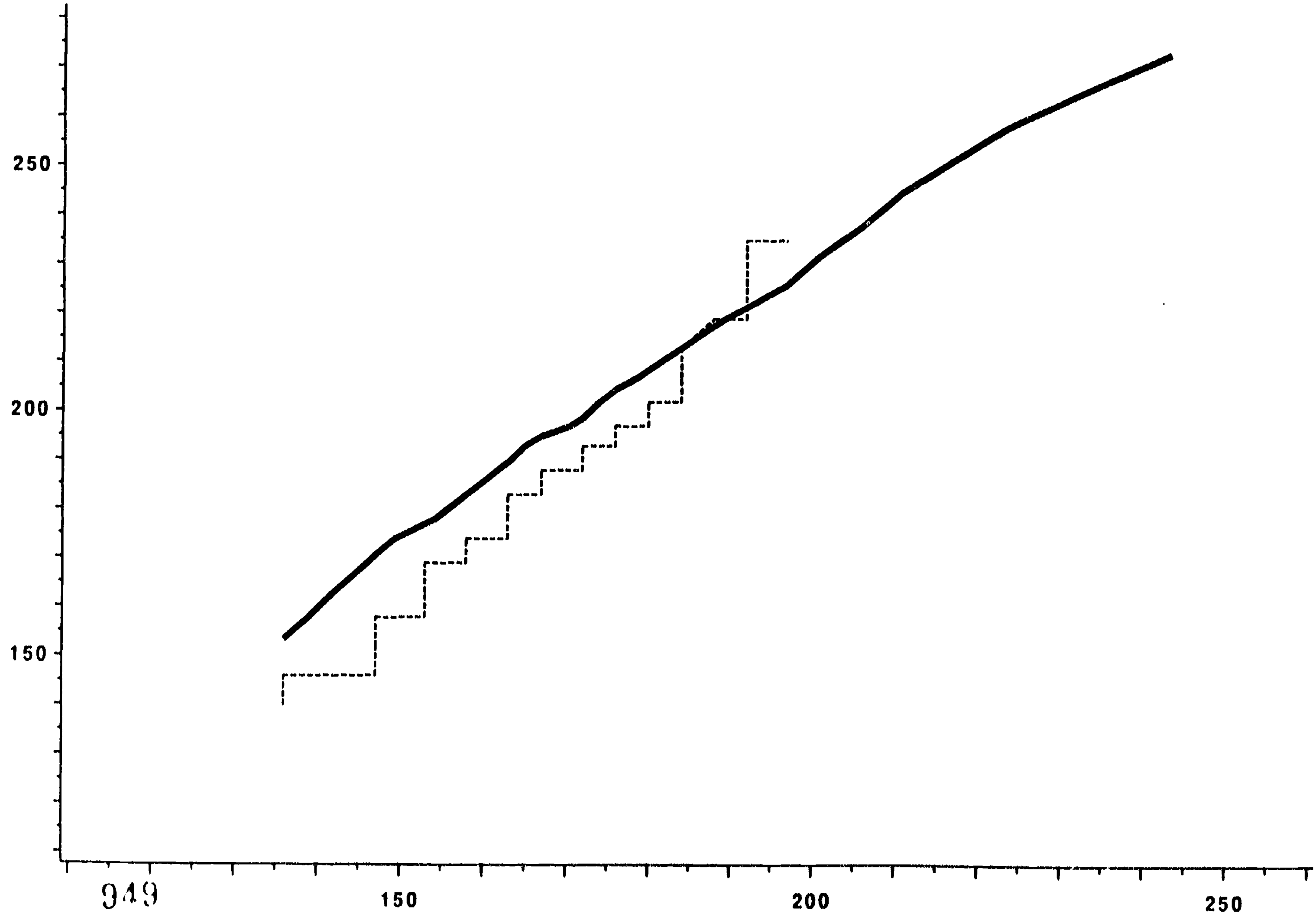
° As a reminder when reading this figure, TAMP

ERIC

Figure 33

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



CURVE ▬▬▬ NORM -------- LE

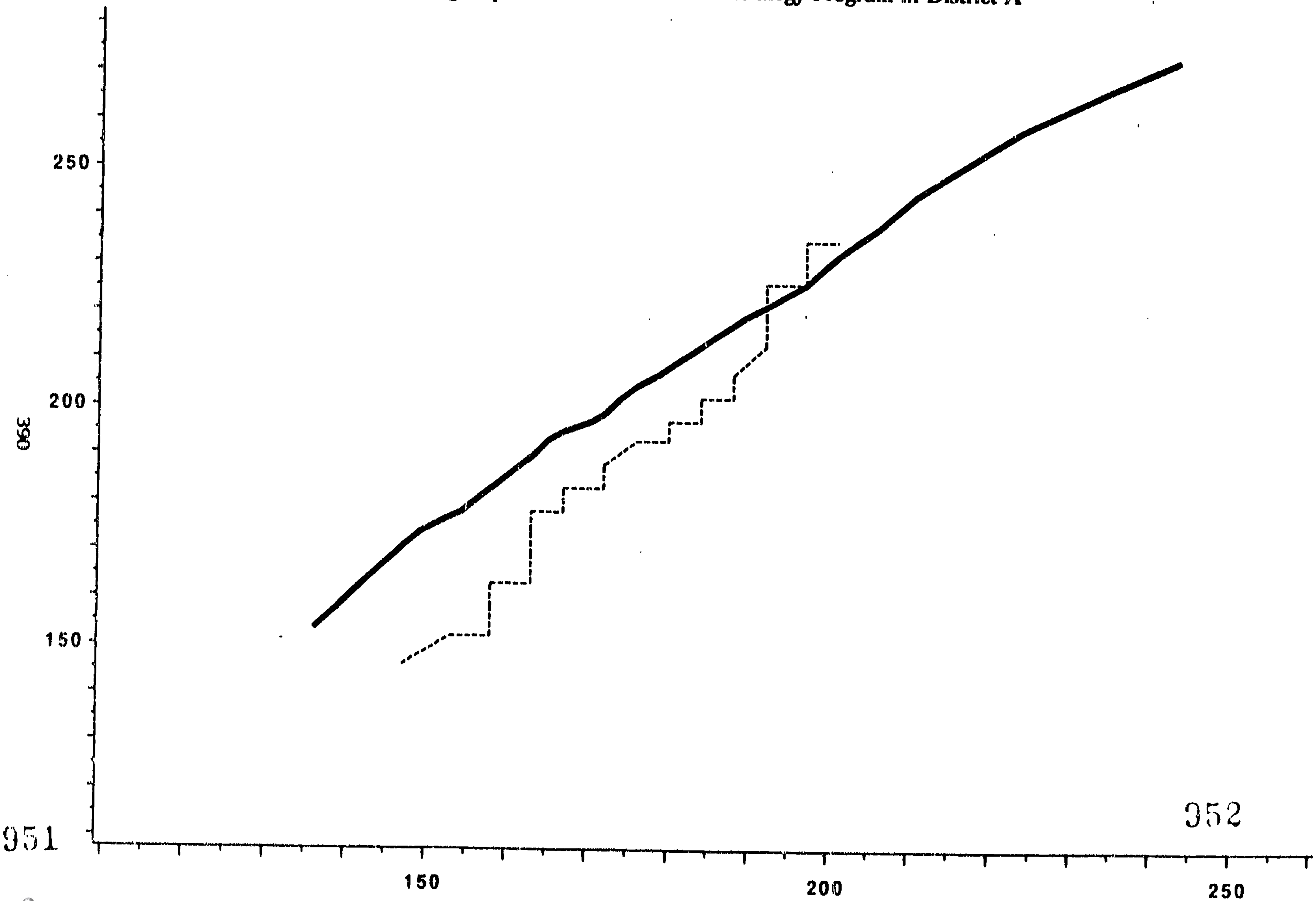
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

389

Figure 34

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



951

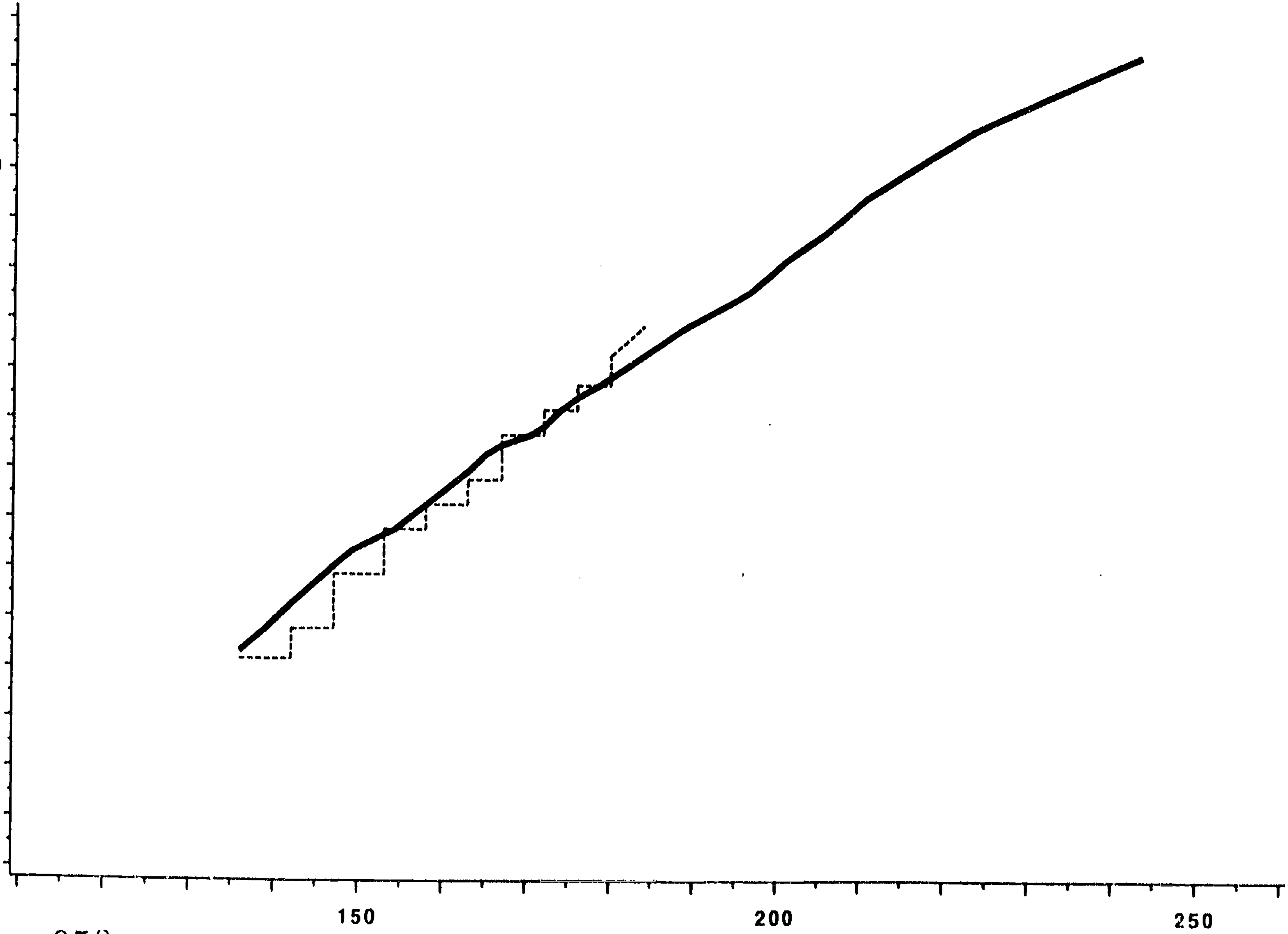
952

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores.

Figure 35

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B

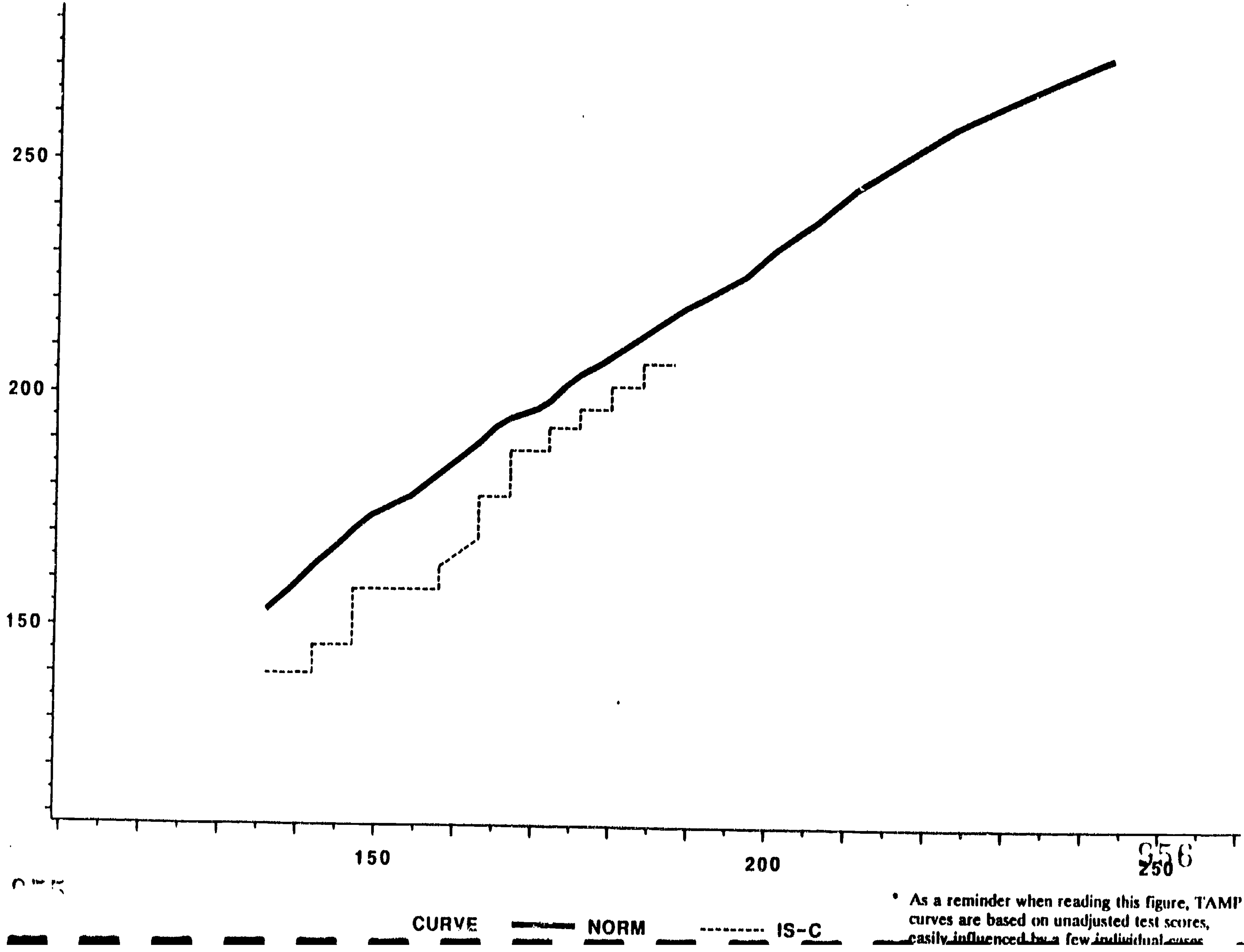


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases and are subject to sampling fluctuation.

Figure 36

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



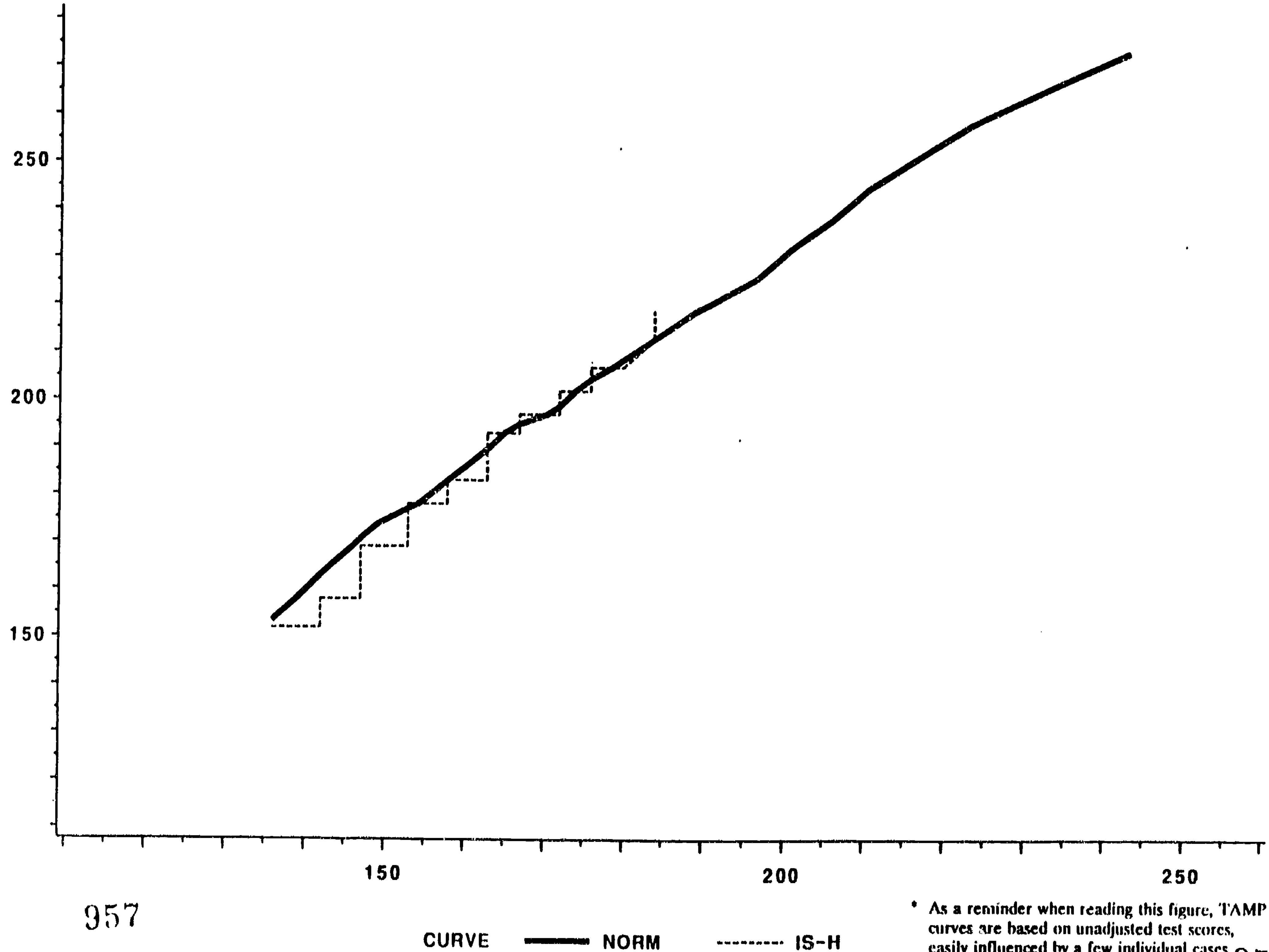
392

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases

Figure 37

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H

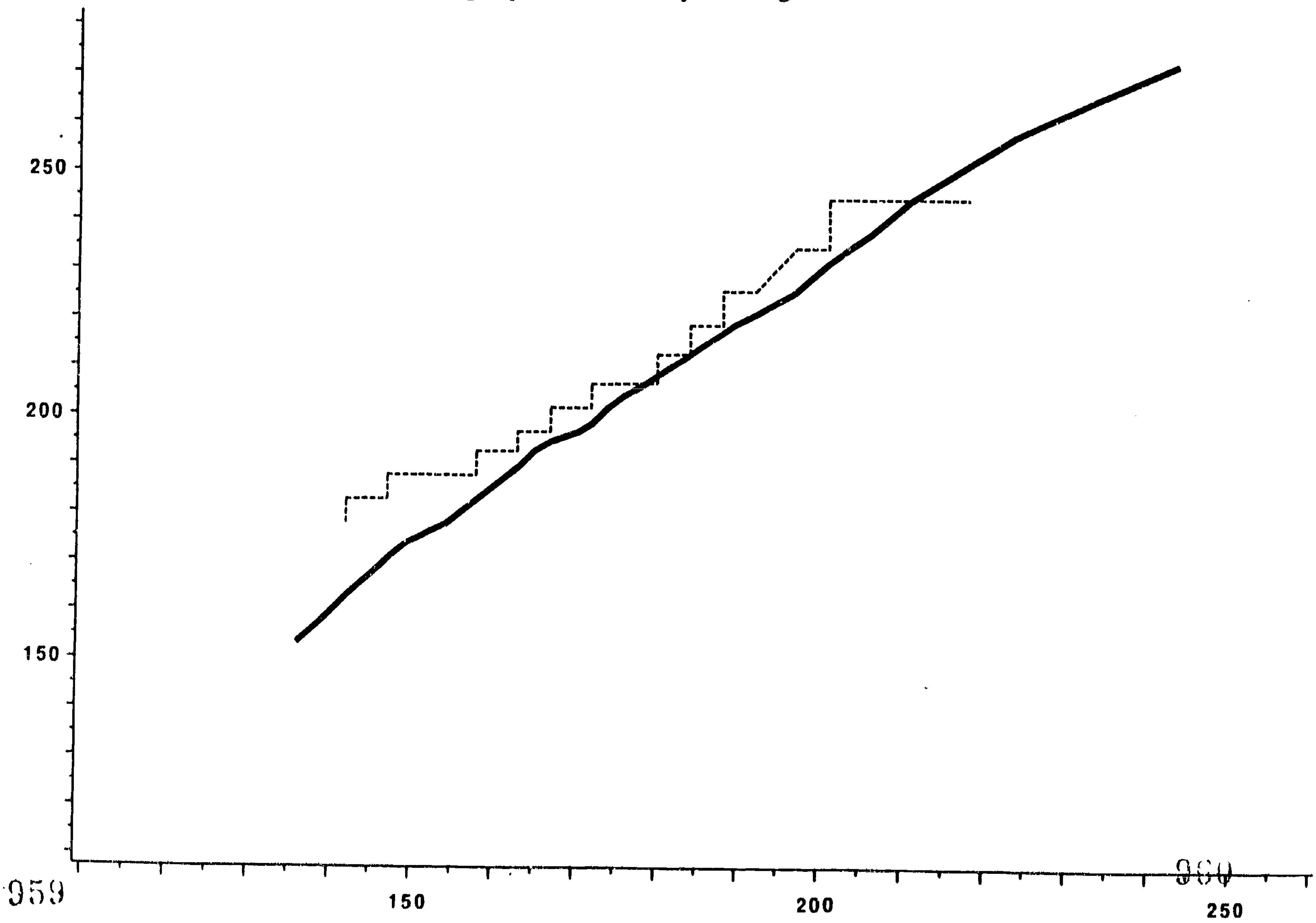


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 38

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



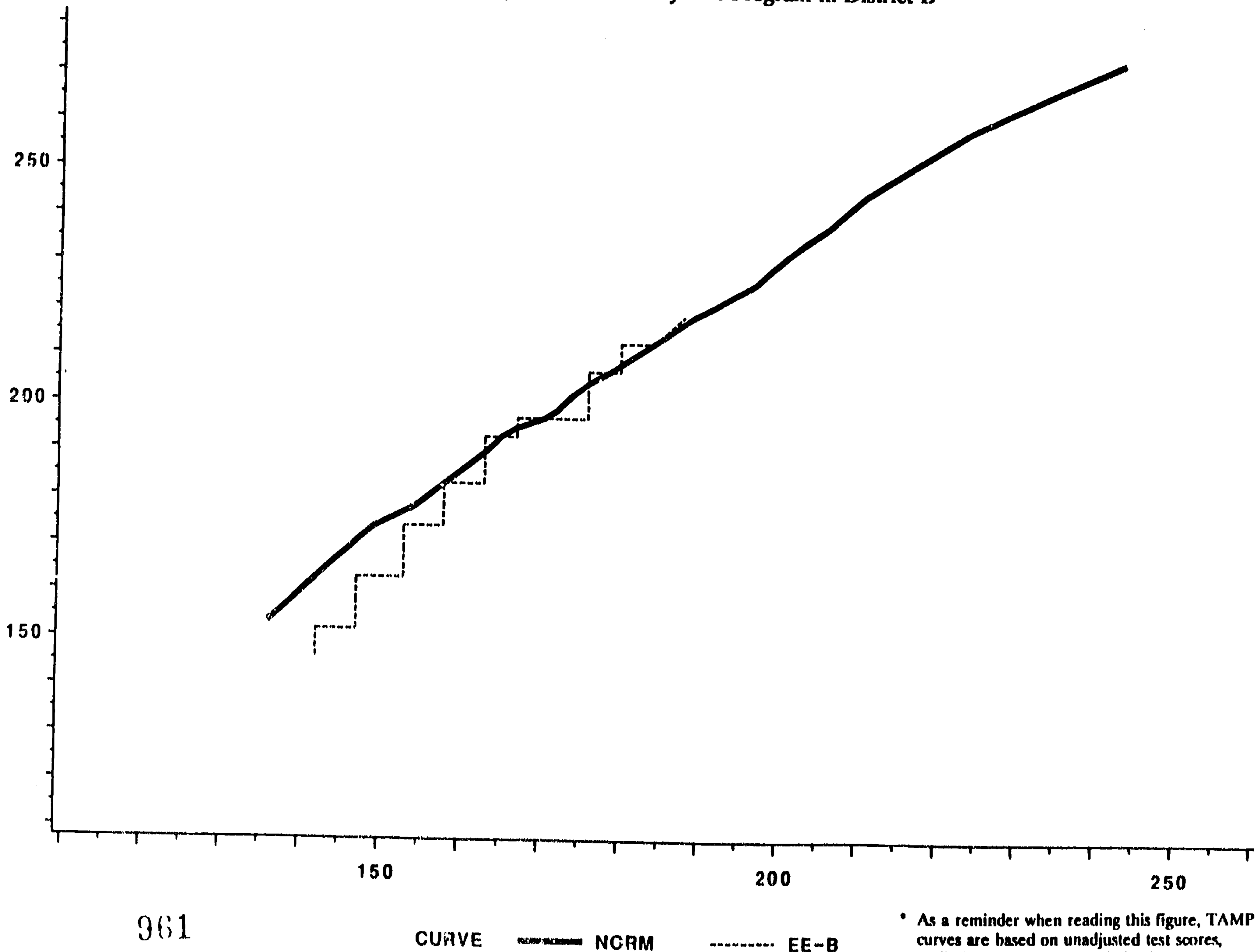
CURVE NORM EE-A

[a] As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 39

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



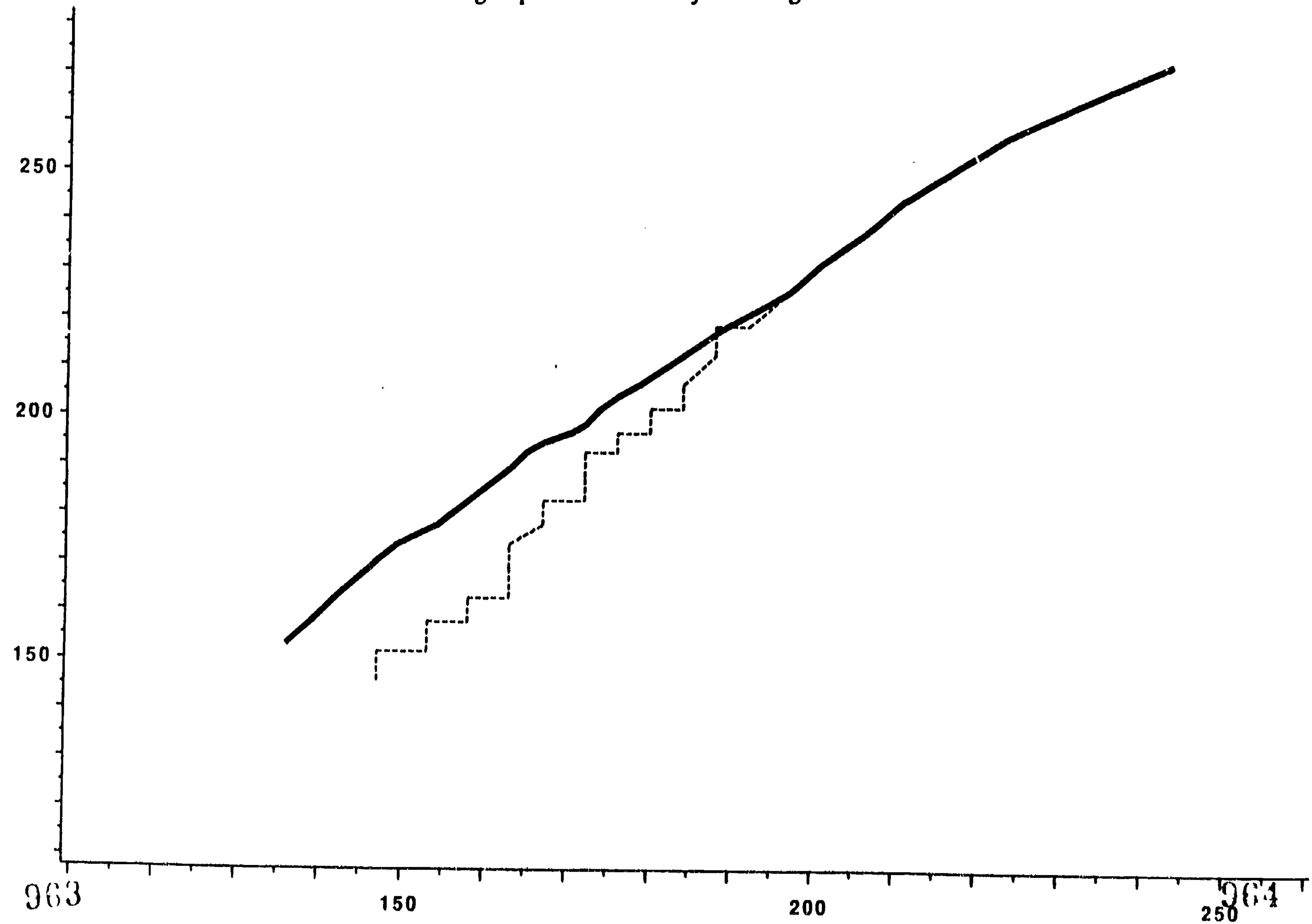
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 40

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C

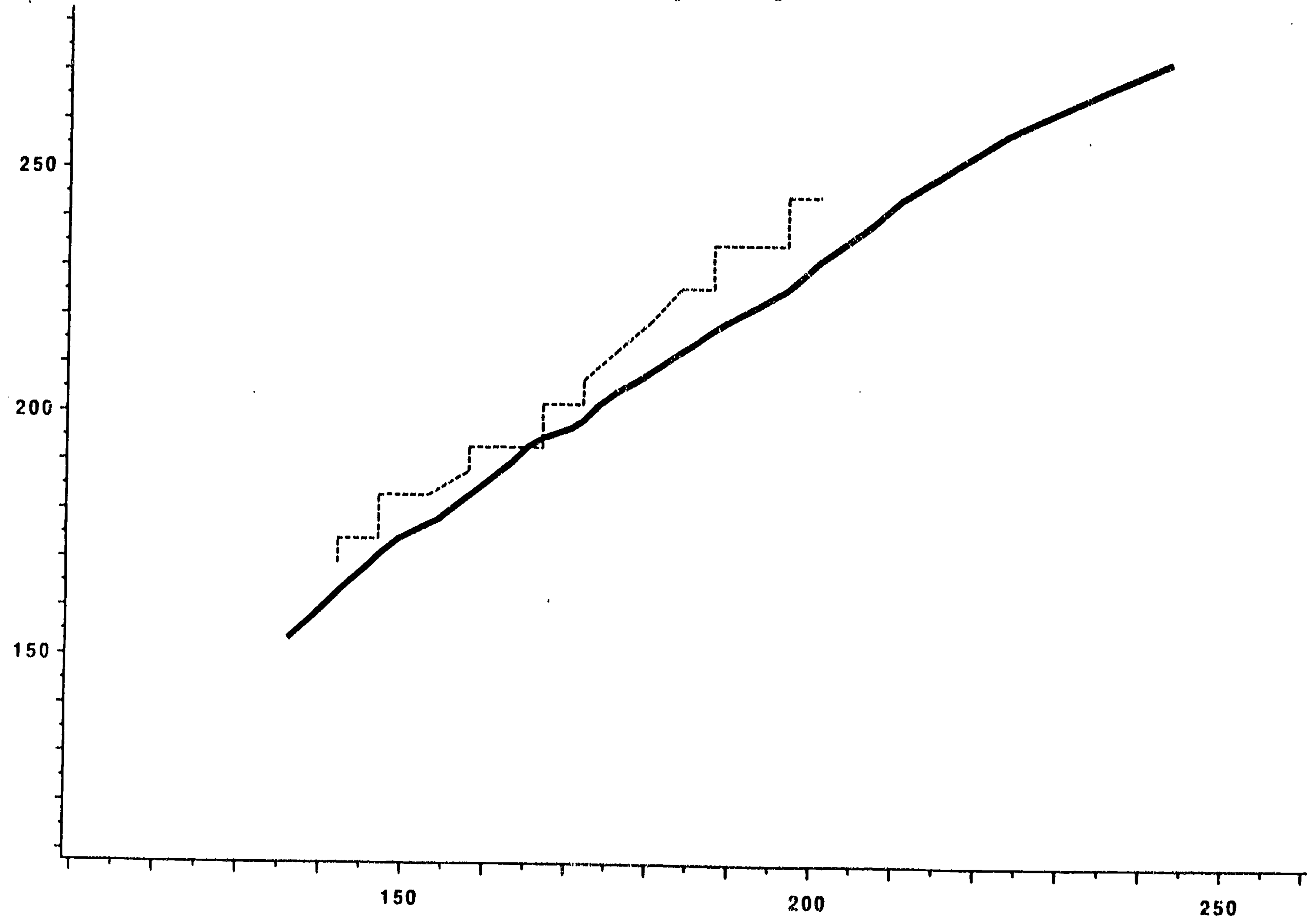


963 964

Figure 41

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

C.	**Grade Span:**	**Kindergarten**	**to**	**Kindergarten**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**Spanish**
	Content:	**Language**	**to**	**Language**

An underlying assumption of primary language programs (i.e., those that use the child's primary language for instruction) is that the more proficient students are in their first language, the faster they will acquire a second language. Thus, it is necessary to examine the extent to which target students across the three instructional programs began and ended kindergarten with comparable primary language skills and how their primary language skills might have improved relative to this norming population.

As they did not receive any instruction in Spanish, it was expected that immersion strategy students would not exhibit improved primary language skills. Consistent with this prediction, it appears that immersion strategy students learned language skills as measured in Spanish slower than this national norming population learned their primary language skills (English) as measured in English (see Figure 42). There was no variation among the individual immersion strategy program sites (see Figures 45 to 48).

Contrary to theoretical predictions, it seems that early-exit students exhibited growth in their Spanish language skills comparable to this norming population and as great as or greater than the growth of late-exit students (see Figures 43 and 44). Minor variation was noted among individual early-exit sites (see Figures 49 to 52), perhaps reflecting the different amounts of primary language instruction provided. The growth rate of early-exit site H students was closest to that of this norming population. Given that the growth rates for the other early-exit sites were slower than for this norming population, the growth rate for site EE-H is most responsible for the composite early-exit TAMP curve approximating the growth rate of this norming population.

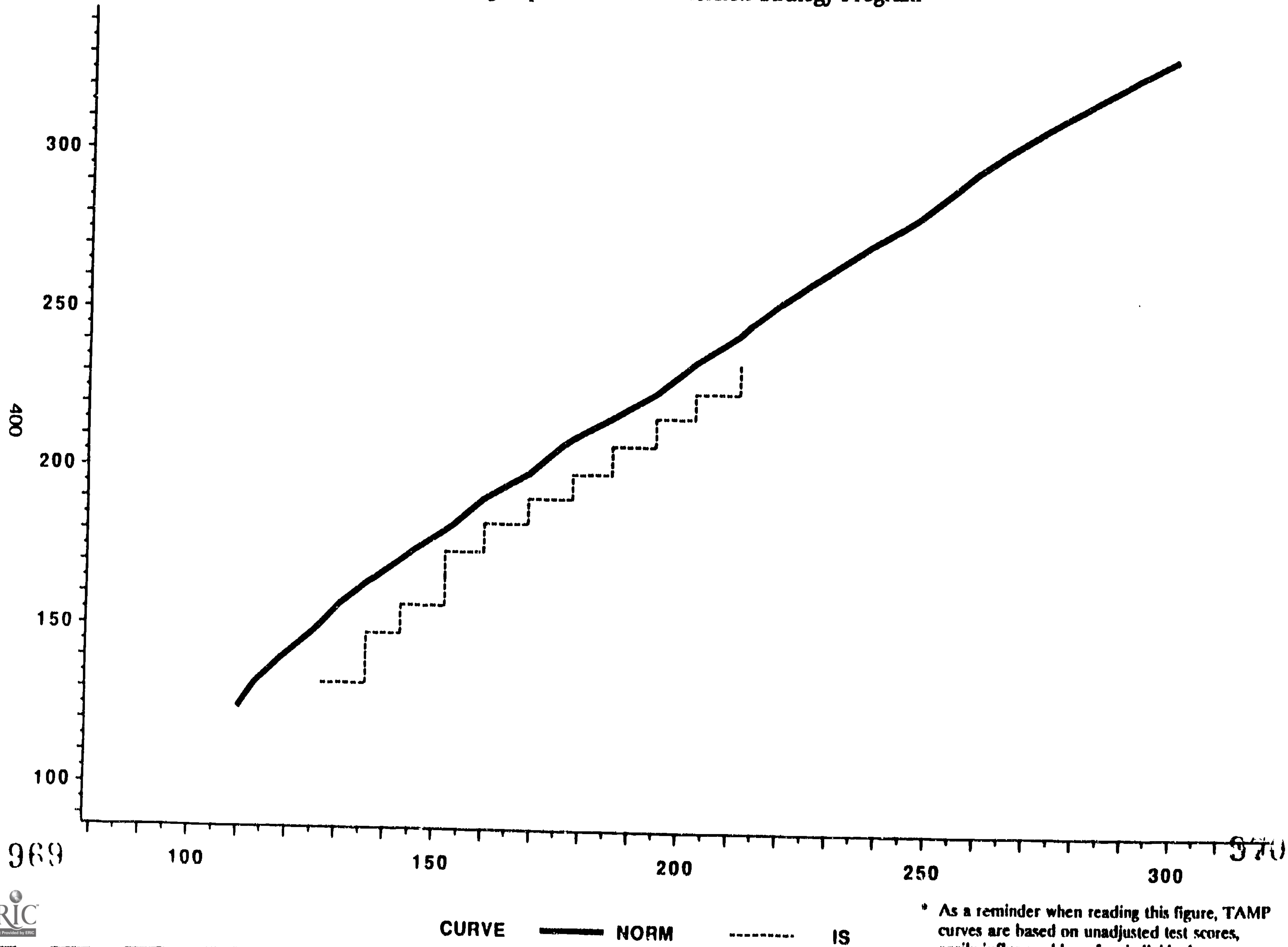
Late-exit students with lower entry-level skills improved in their Spanish language skills at about the same rate as did this norming population (see Figure 44). There is some indication that late-exit students with higher entry-level Spanish language skills may have lost ground relative to comparable students in this norming population and in the early-exit program. The latter indicates that the curricular and instructional efforts of the late-exit program might not have adequately addressed the needs of these students, and suggests that curriculum and instruction need to be expanded to better accommodate the learning needs of students with higher entry-level Spanish skills.

In sum, relative to this norming population, target students in all three programs appear to have begun kindergarten with lower primary language skills and to have learned almost as fast as or slightly slower than this norming population with respect to Spanish language skills. Consistent with predictions, immersion strategy students as a group lost the most ground relative to this norming population. Contrary to expectations, early-exit students, although growing more slowly relative to this norming population, acquired primary language skills at a slightly higher rate than did the immersion strategy students; no difference had been expected.

Figure 42

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

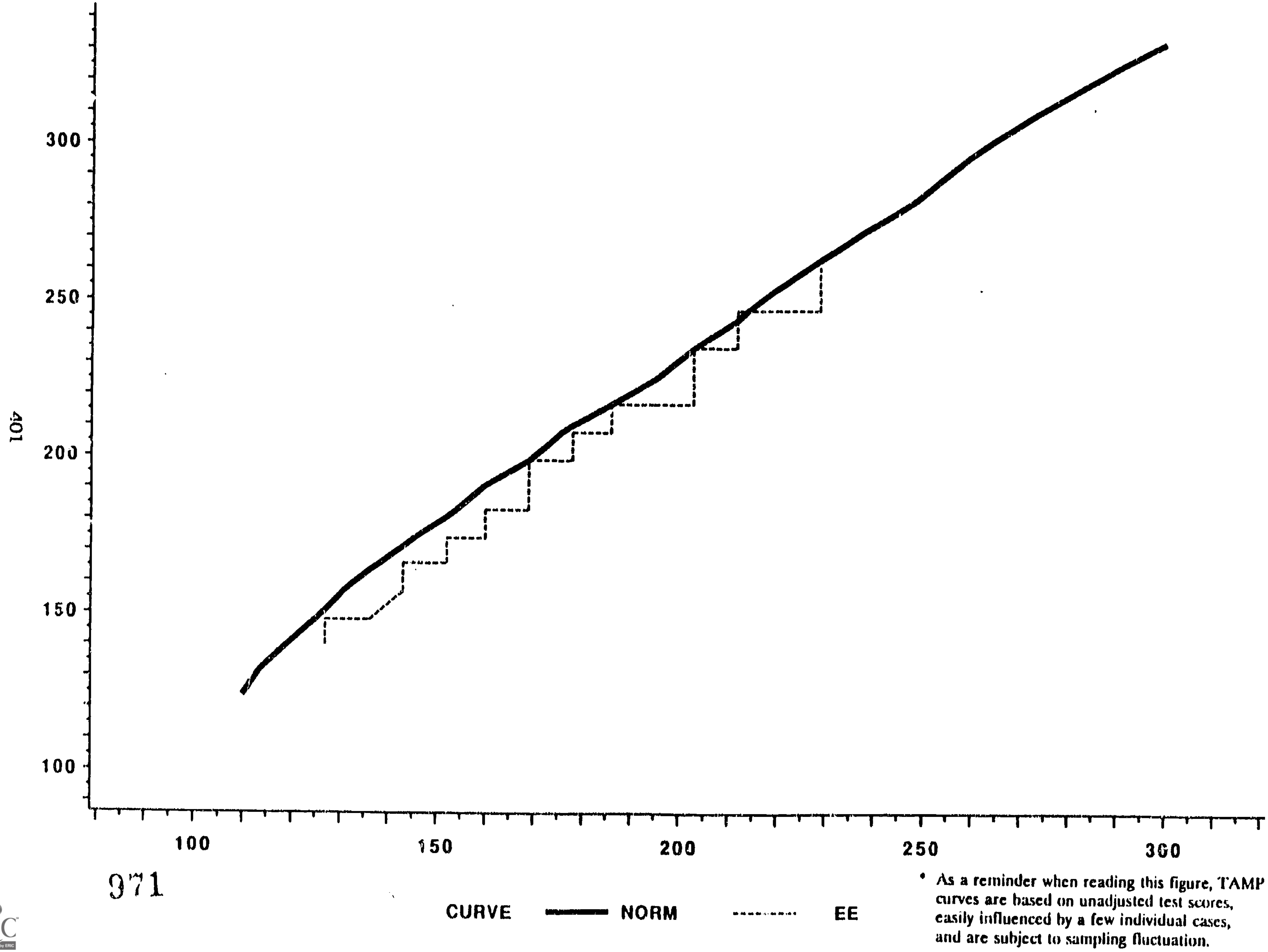


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 43

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



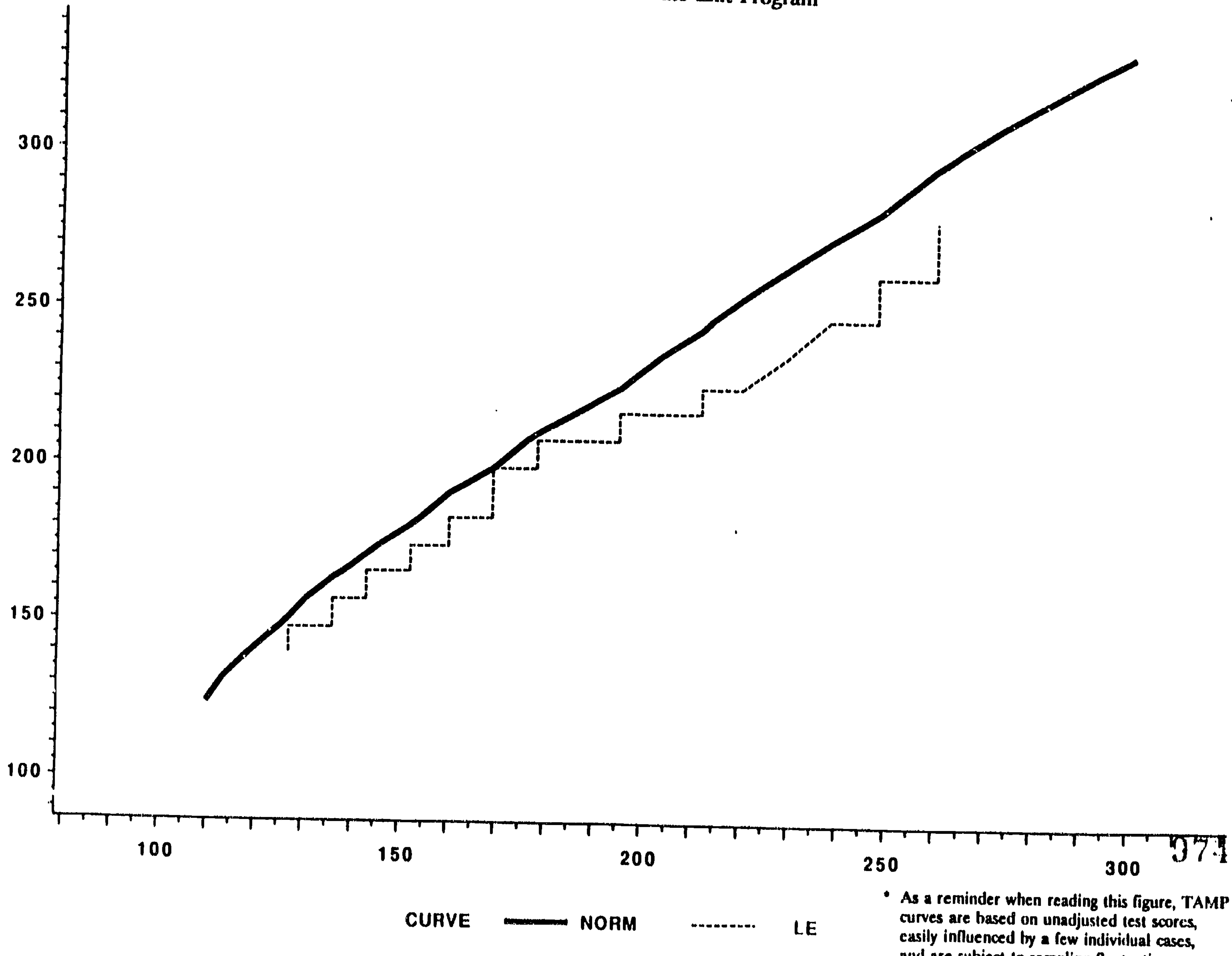
971

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 44

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



402

973

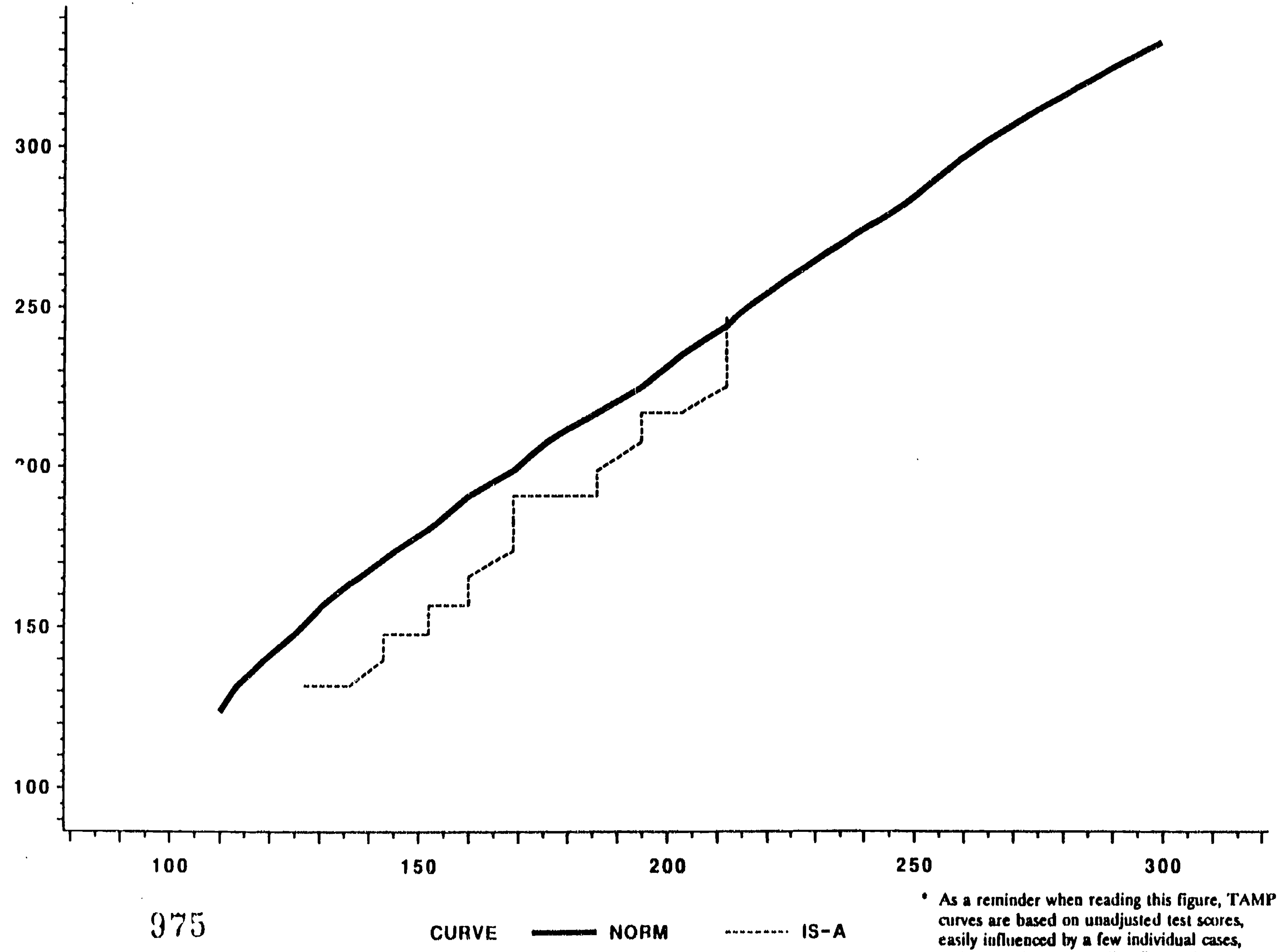
974

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 45

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



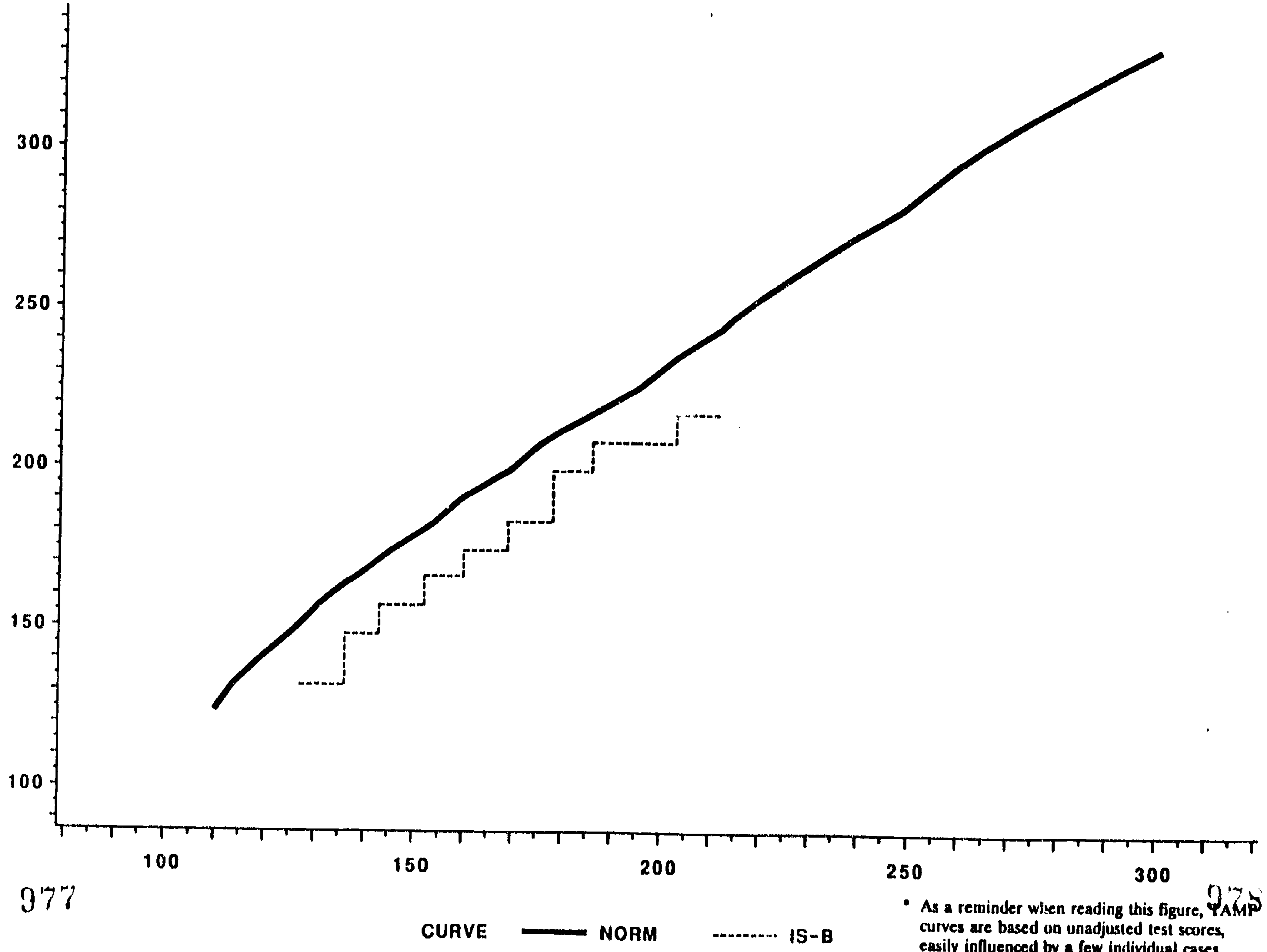
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 46

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



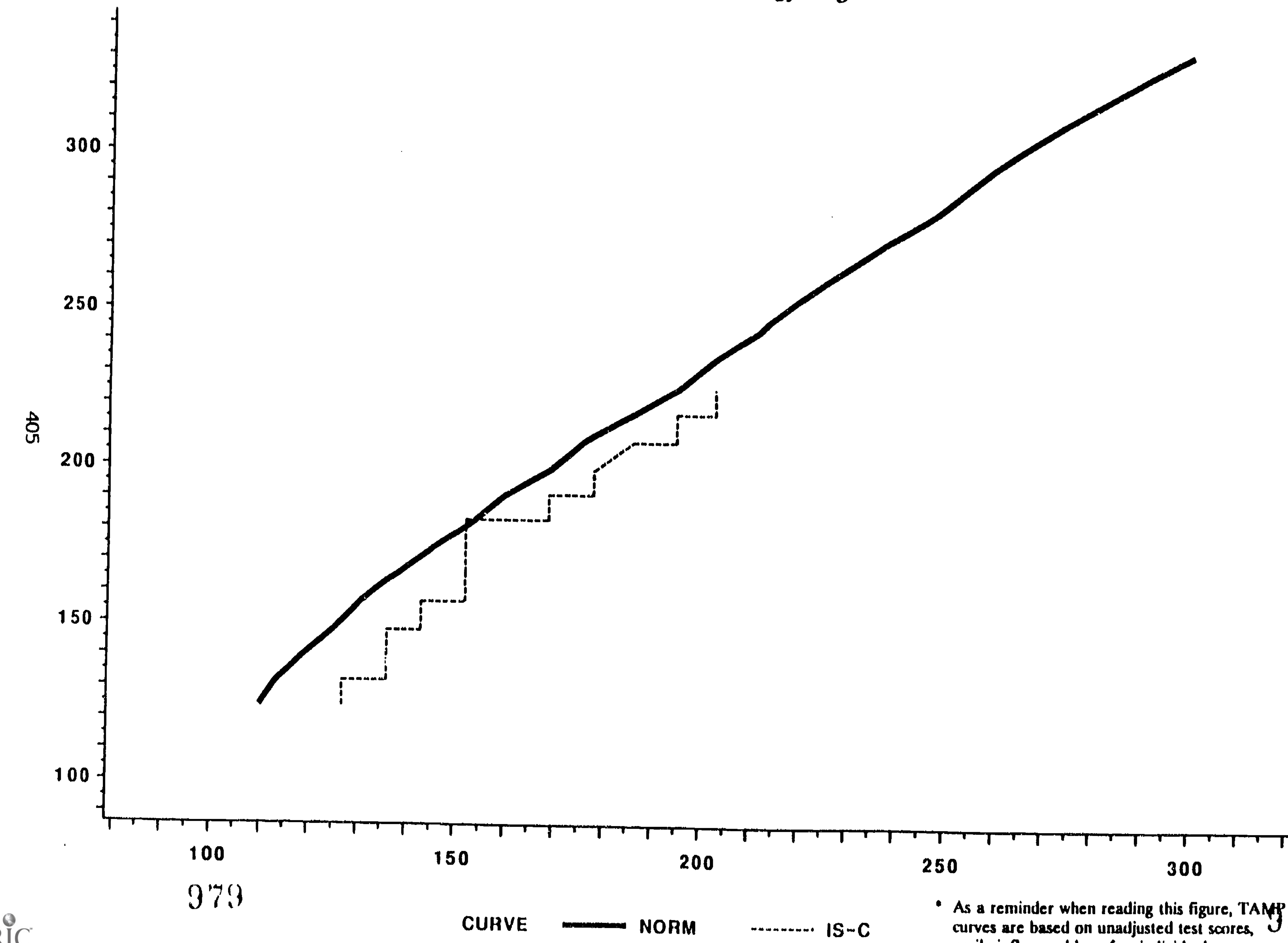
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 47

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C

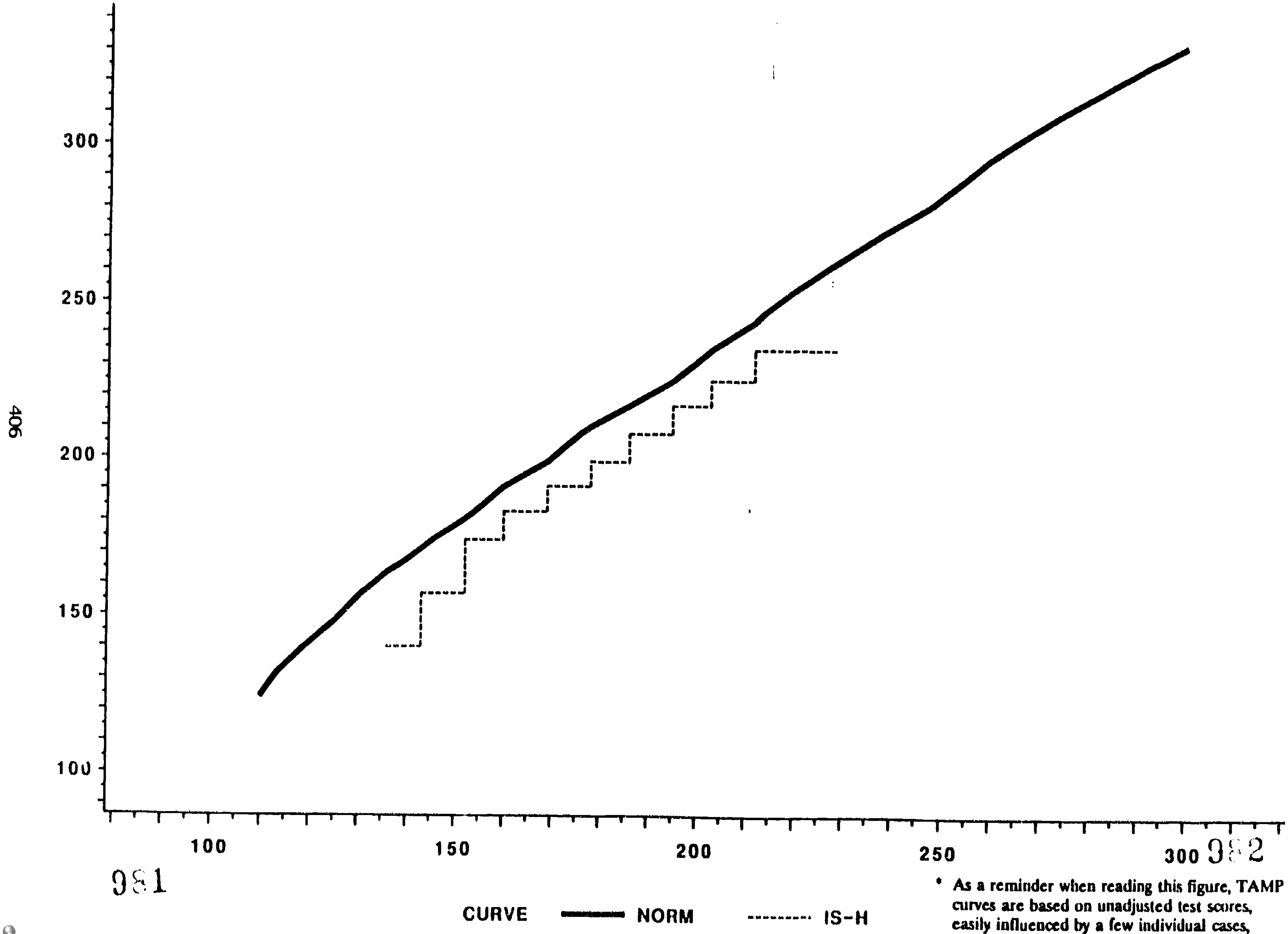


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 48

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H



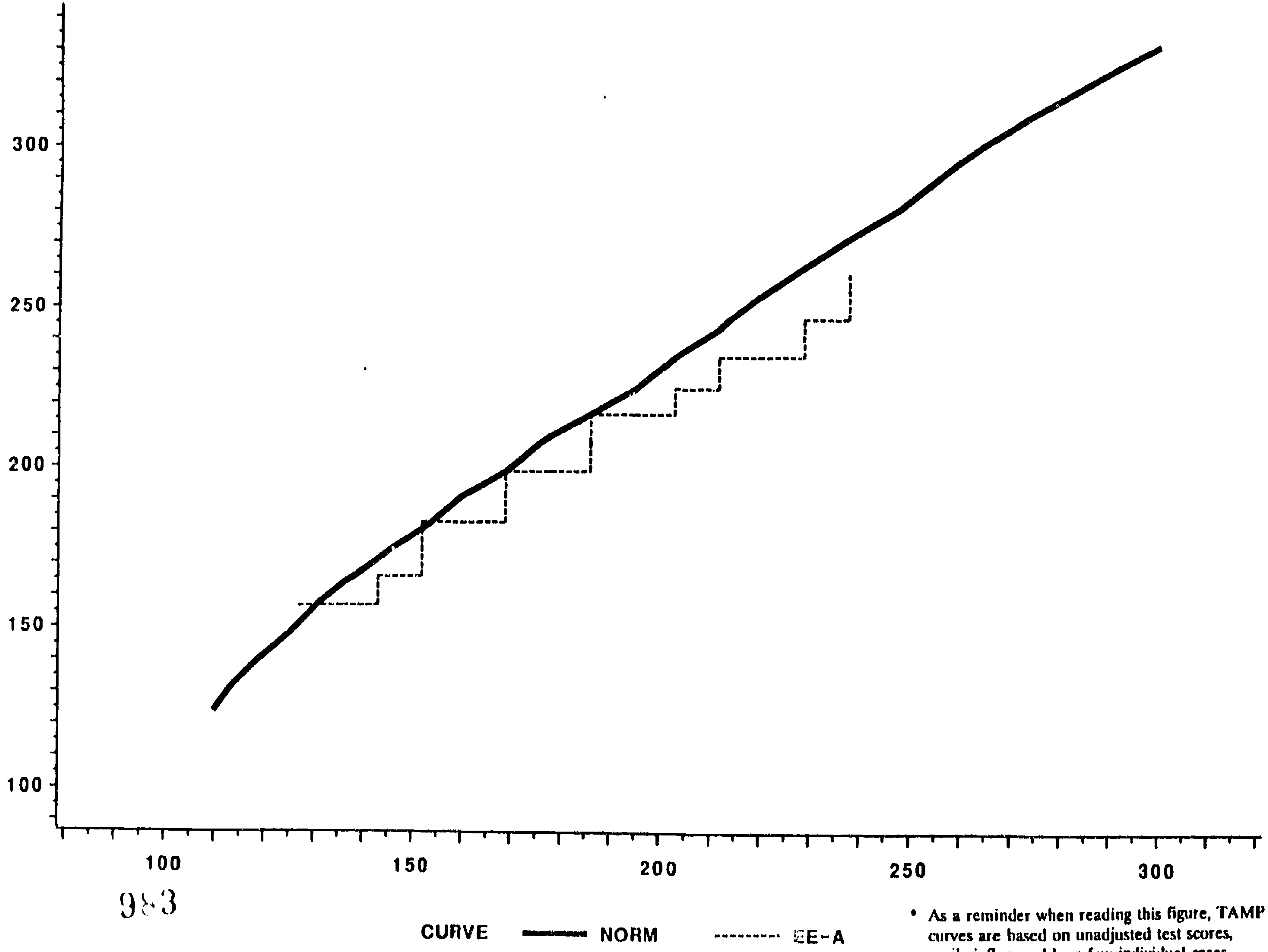
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC Full Text Provided by ERIC

Figure 49

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



407

983

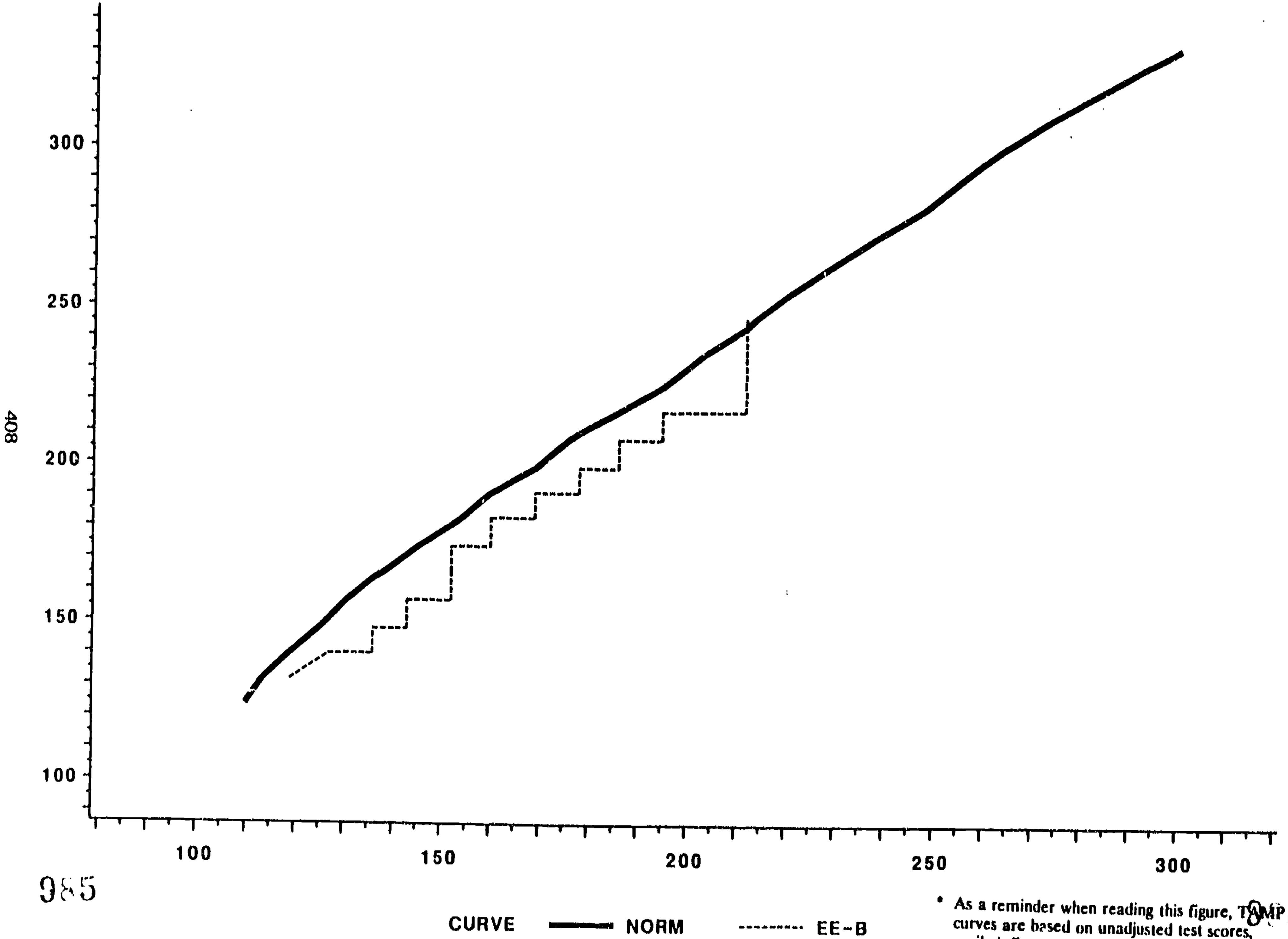
984

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 50

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B

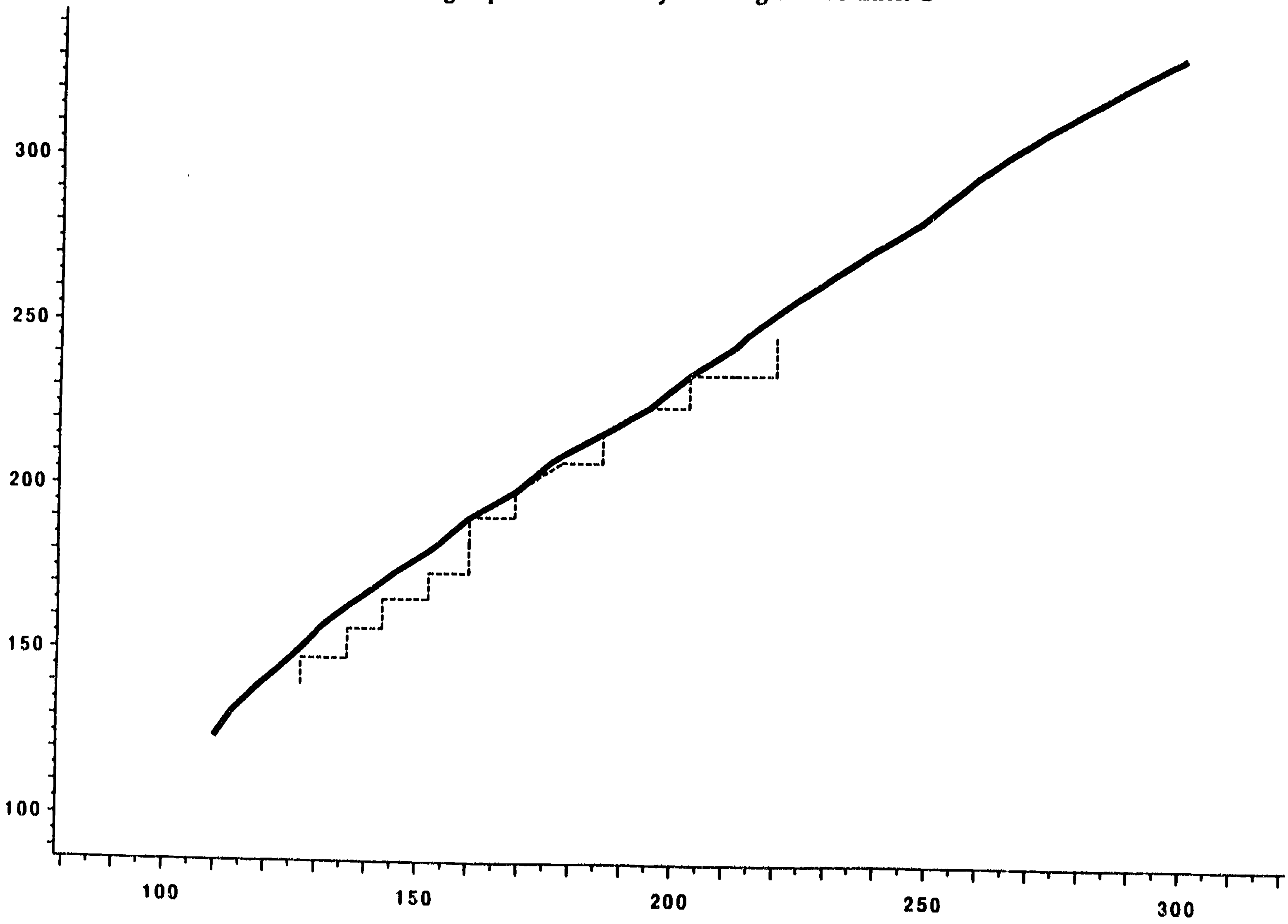


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 51

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



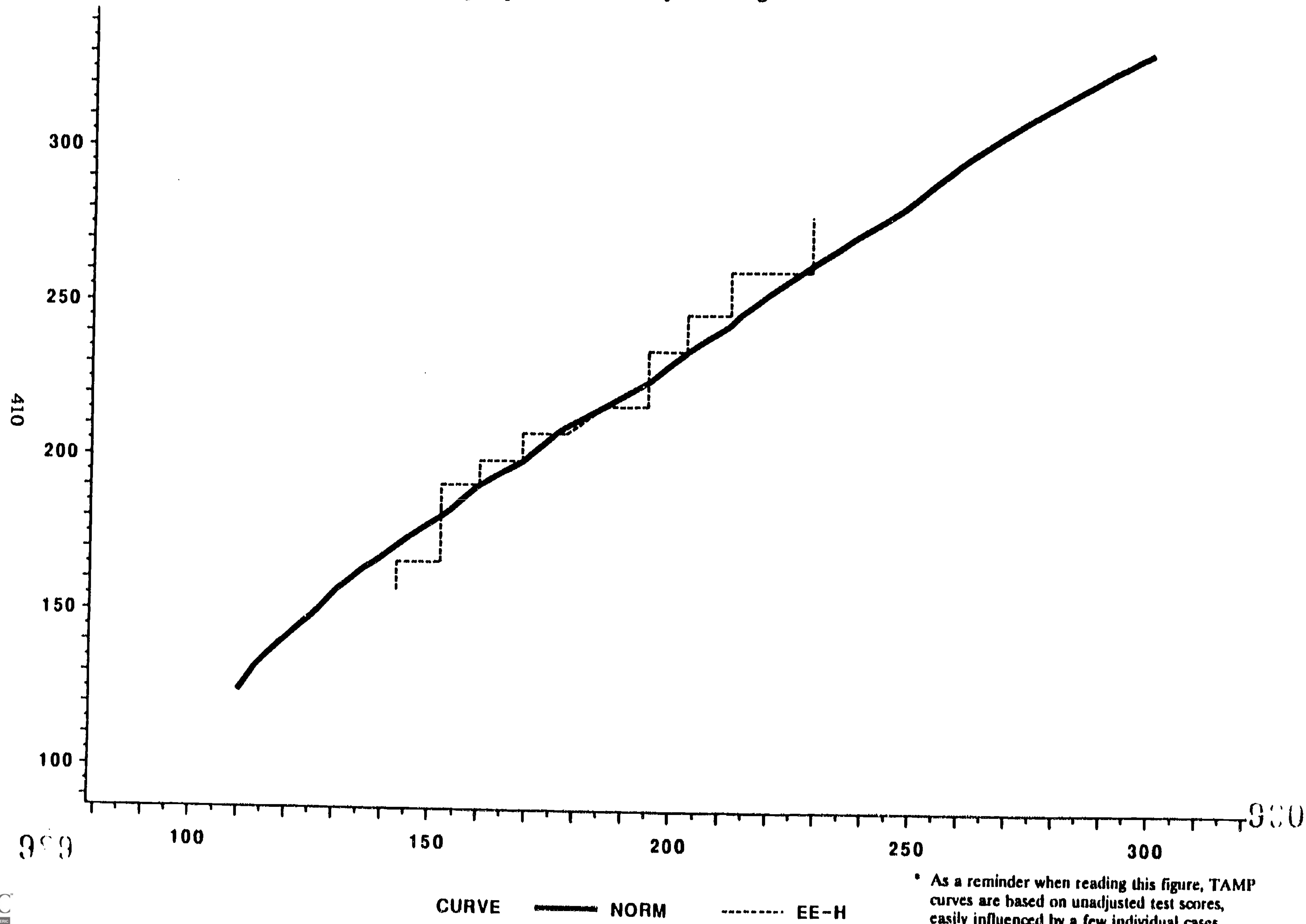
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 52

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

D.	**Grade Span:**	**Kindergarten**	**to**	**Kindergarten**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Language**	**to**	**Language**

Figures 53, 54, and 55 show that target students in each program began their kindergarten year within approximately the same range of Spanish language skills, which was slightly lower than that of this norming population. Student growth rates in each program over the year when tested in English language skills appear to be somewhat similar; these students generally seemed to grow more slowly than this norming population. The upper end of the distribution of scores for immersion strategy and early-exit students is close to these norms; however, the late-exit students definitely had lower growth than these norms. Figures 56 through 59 show that there is almost no variation between individual immersion strategy sites. There is minor variation between early-exit sites, as site EE-A students grew as fast as this norming population, and sites EE-B and EE-C grew slower than this norming population (see Figures 60 to 62). Early-exit site H students with the lowest entry-level skills appeared to grow slower than this norming population, while those with the highest initial scores seemed to grow faster than this norming population (see Figure 63). These patterns suggest that the growth rates in English language skills among immersion strategy and early-exit students were similar to each other and seemed to be generally higher than for late-exit students.

The patterns in student growth seemed consistent with the predictions for each of the instructional models. As students in the immersion strategy and early-exit program received more instruction in English language arts than late-exit students, the former were expected to achieve more than late-exit students by the end of kindergarten. Nonetheless, the difference appears to be slight among the students starting relatively low, indicating that low-scoring students in all three programs ended kindergarten with approximately the same English language arts skills. Among the students scoring above about 200 on the fall kindergarten

ERIC

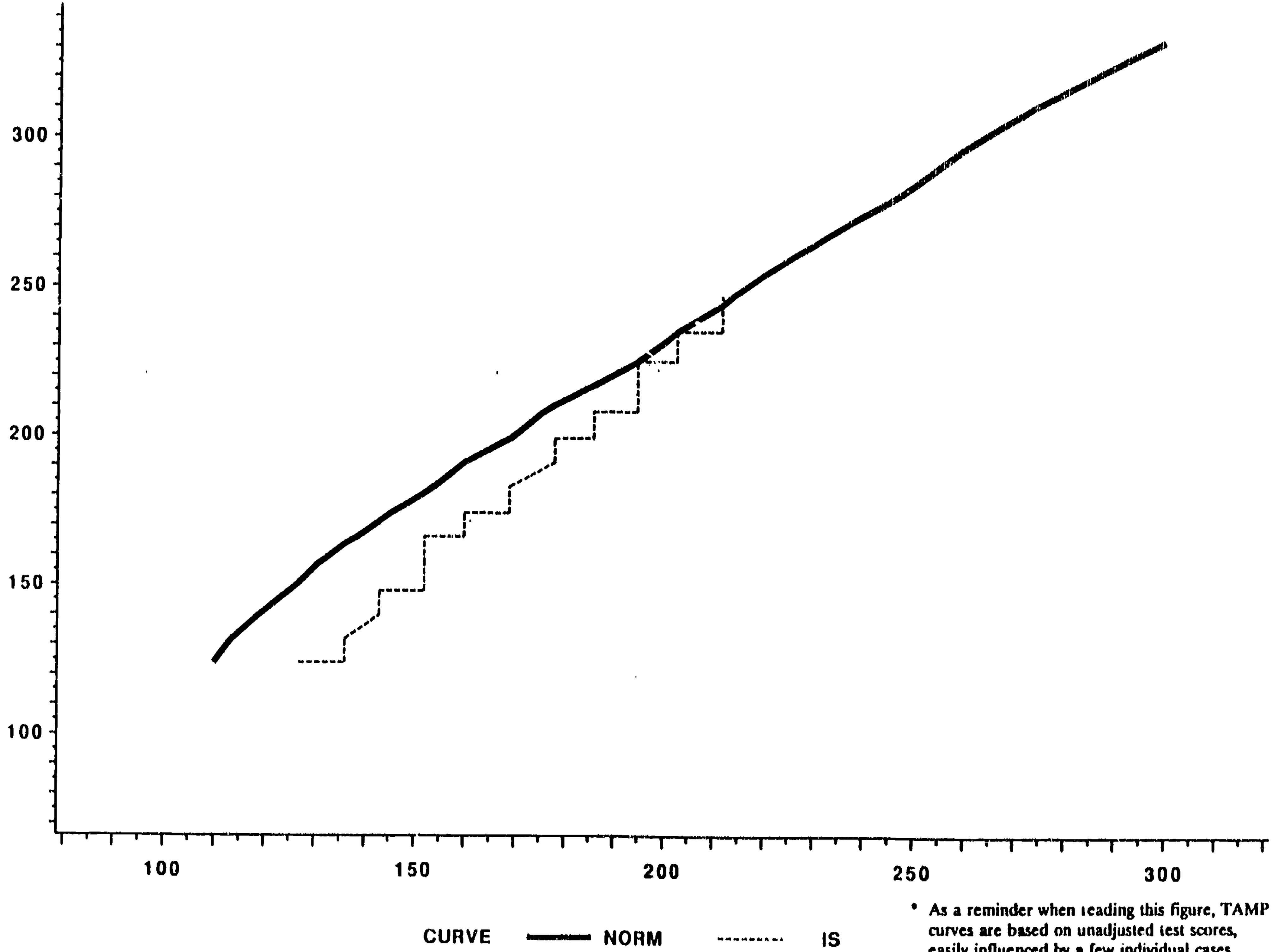
Spanish language test, however, there seems to be a substantial gap between IS/EE and LE. The IS/EE students in this range appear to have grown about as much as was expected from the national norms, while comparable LE students seemed to be losing ground.

ERIC

Figure 53

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

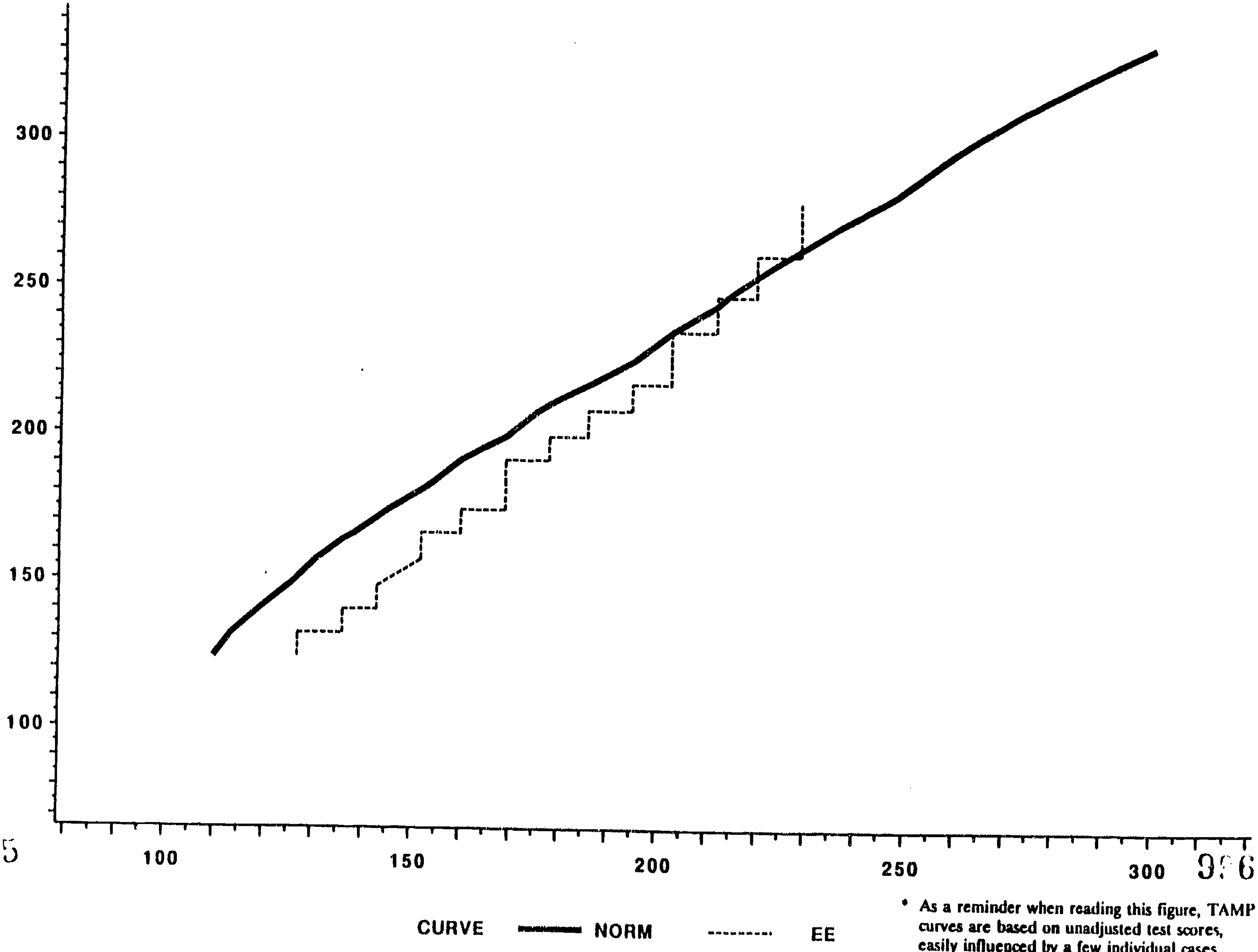


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 54

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



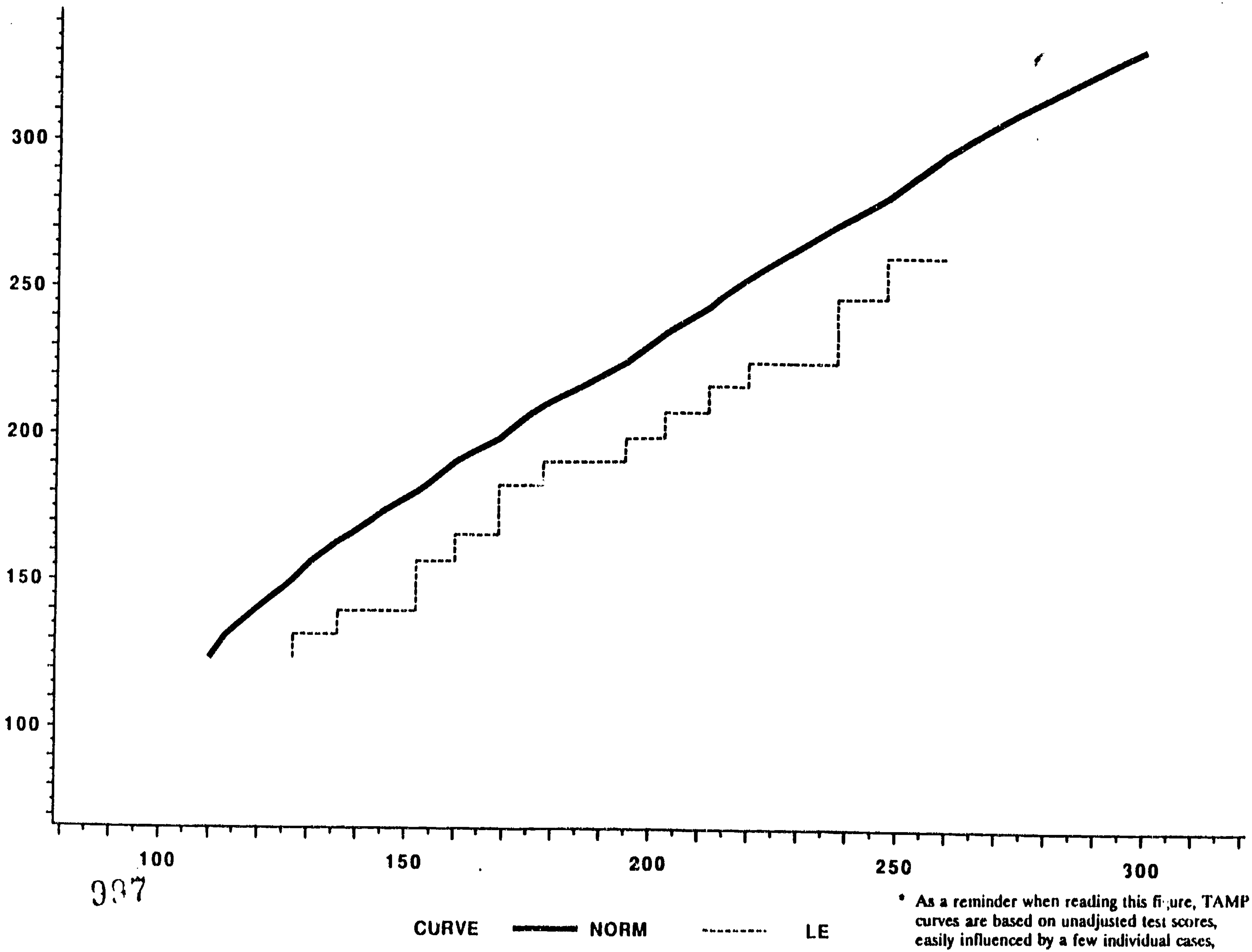
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 55

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program

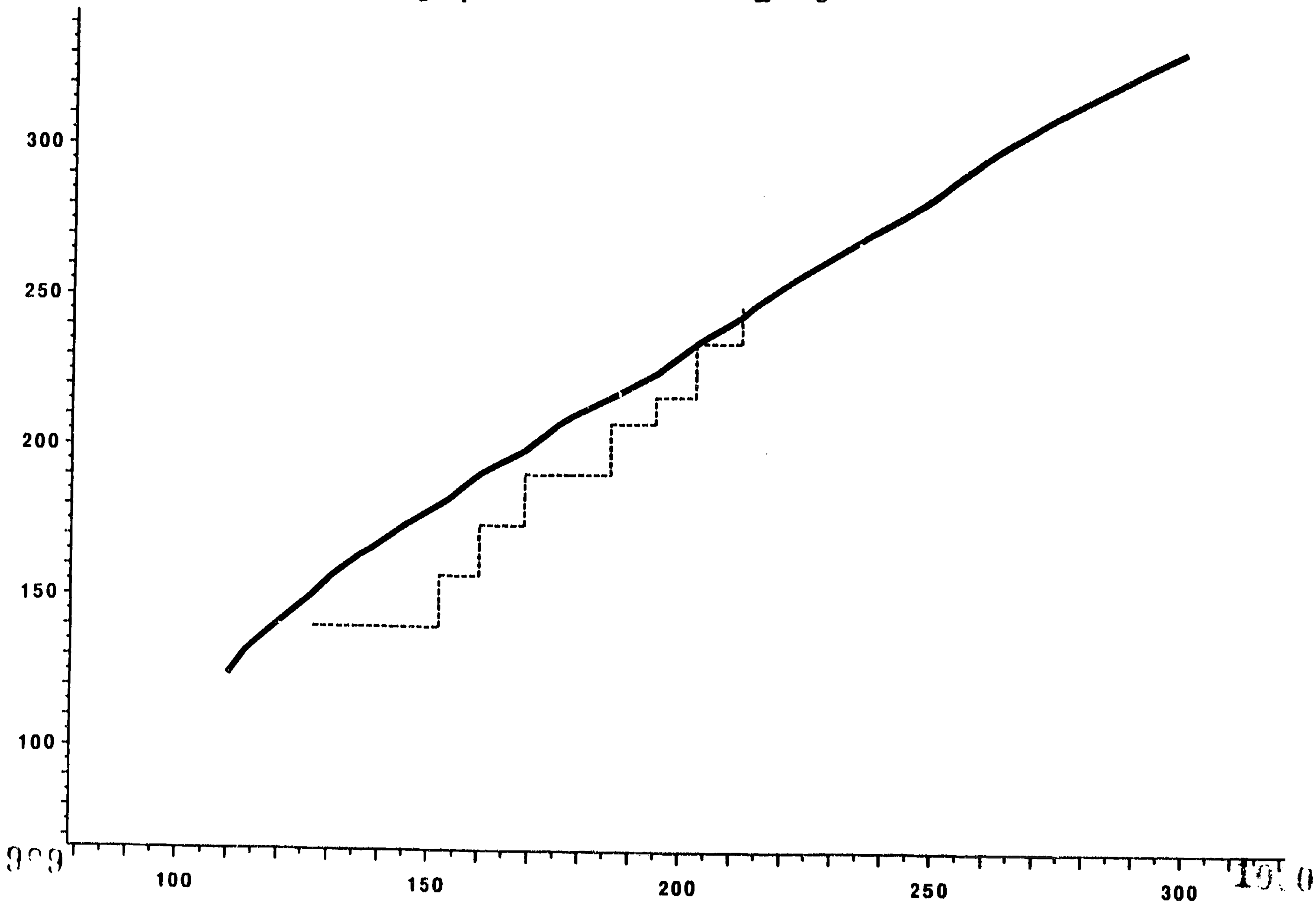


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 56

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



416

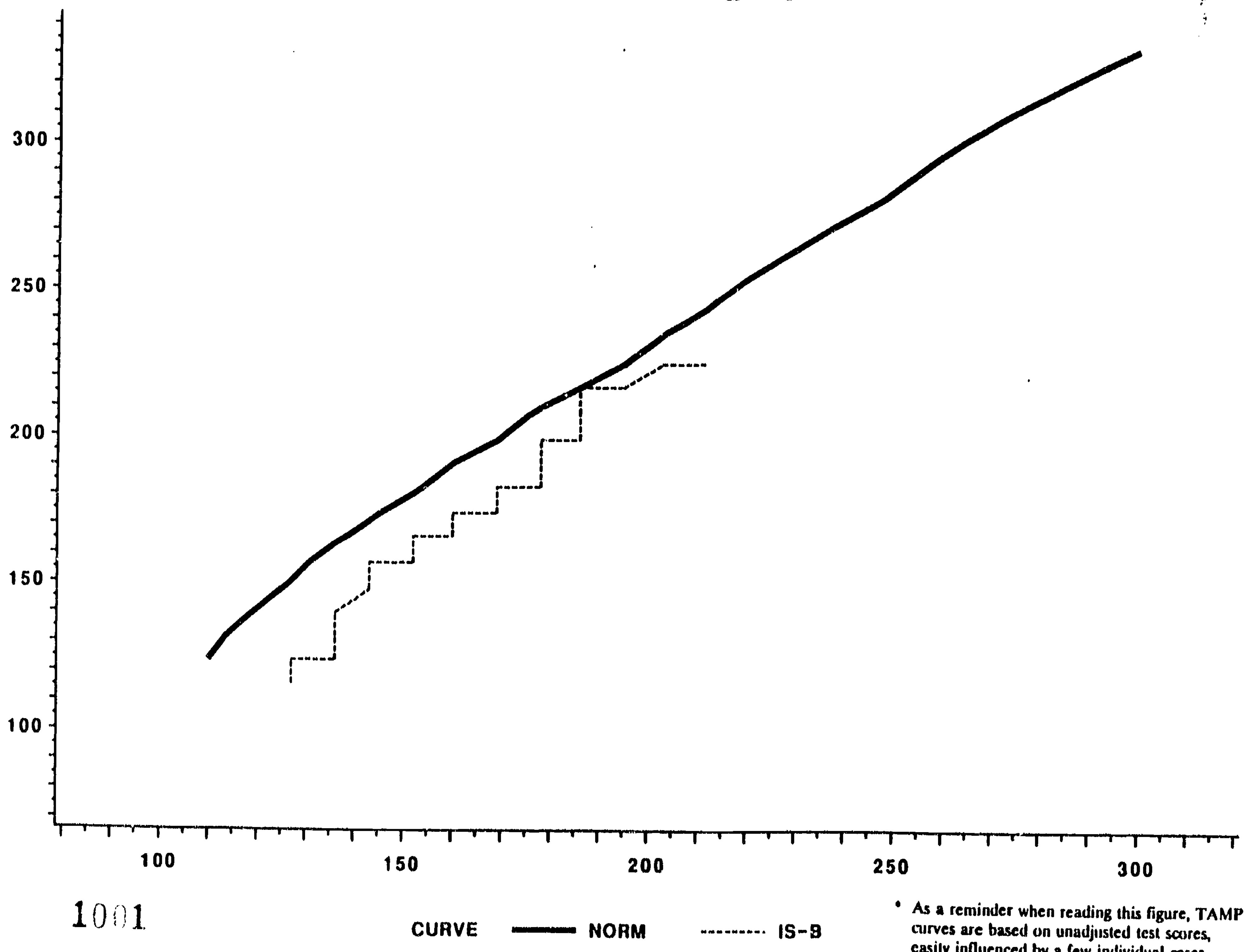
- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 57

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



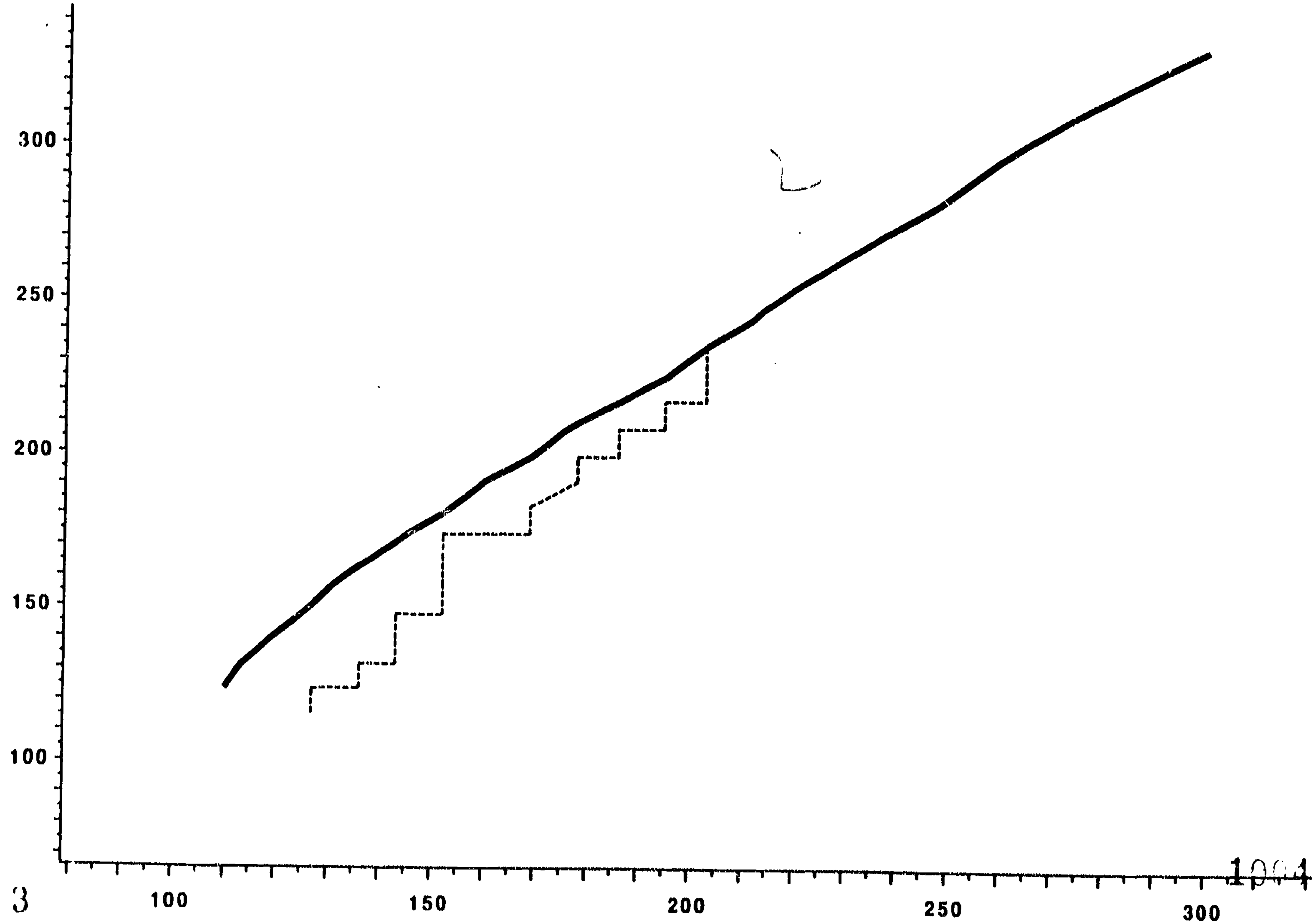
417

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 58

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



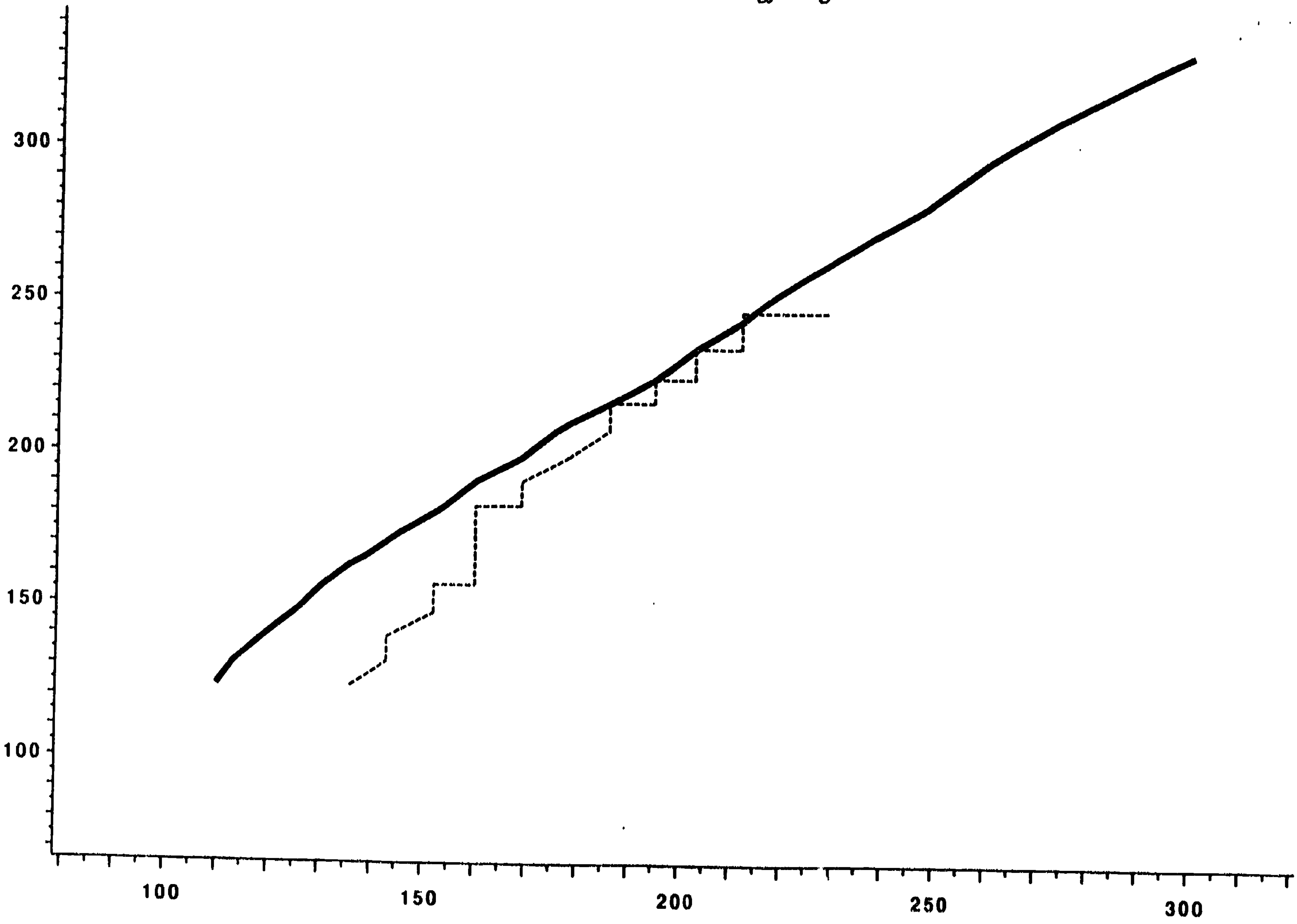
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores.

ERIC

Figure 59

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H



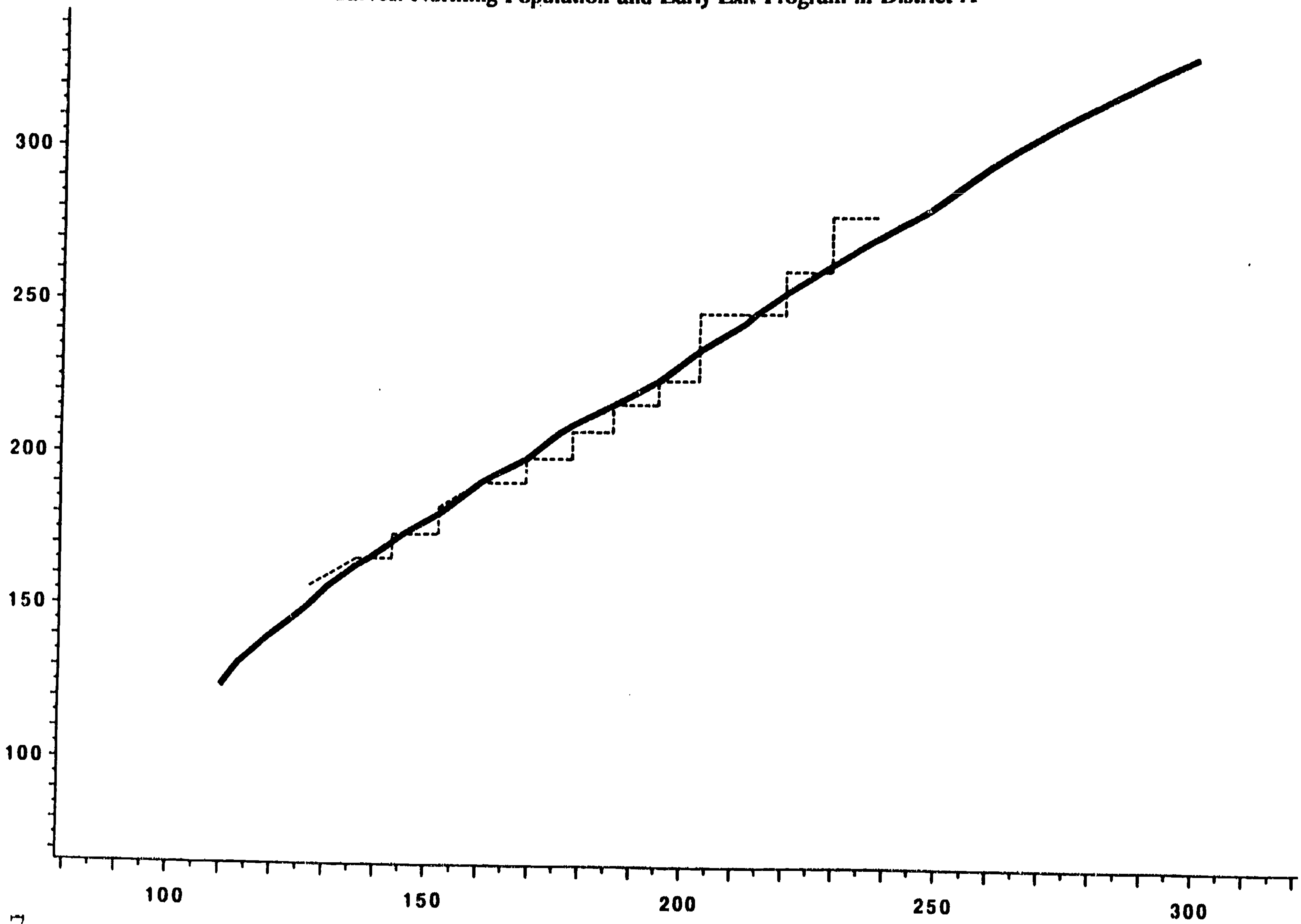
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 60

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



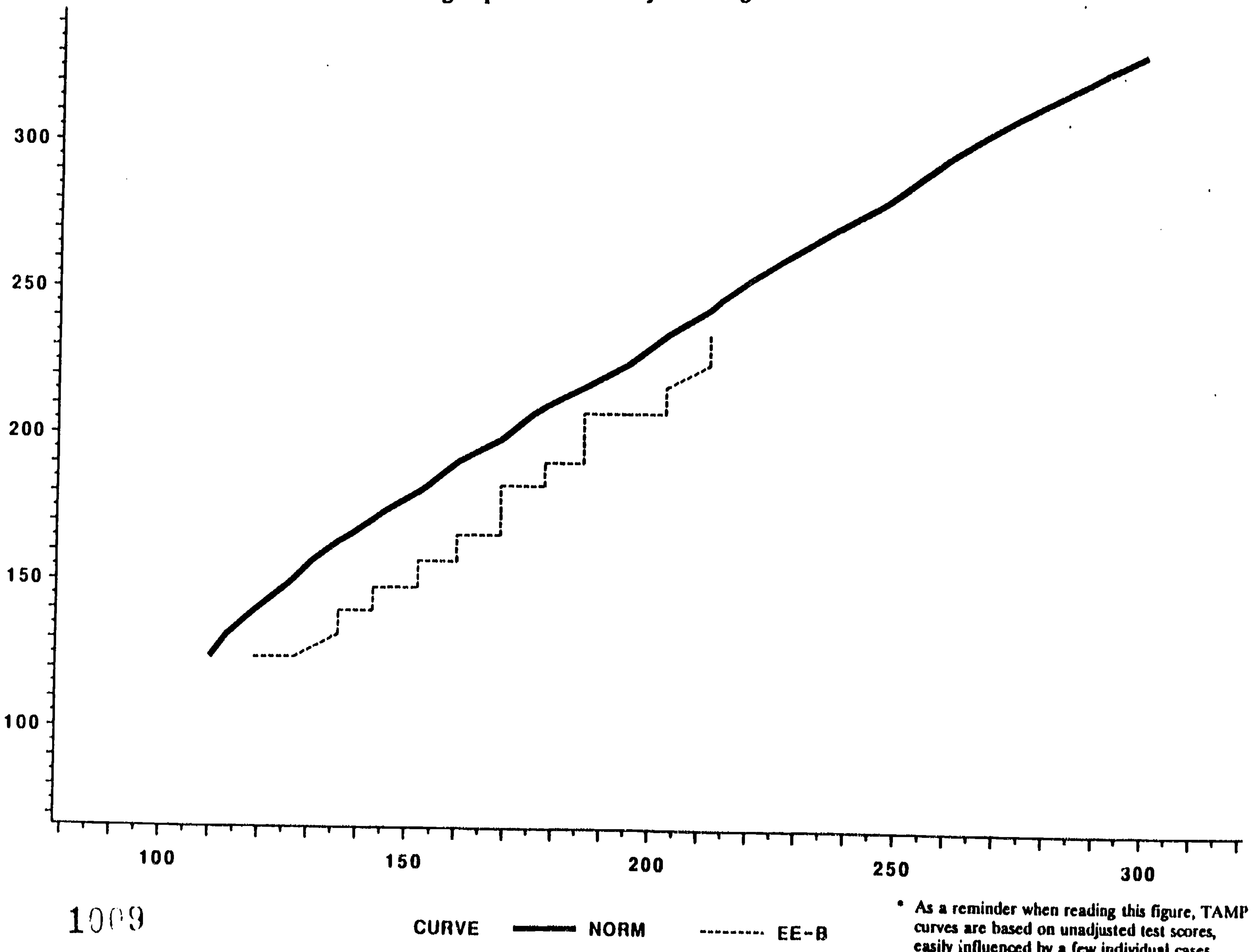
CURVE —— NORM - - - - EE-A

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 61

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



421

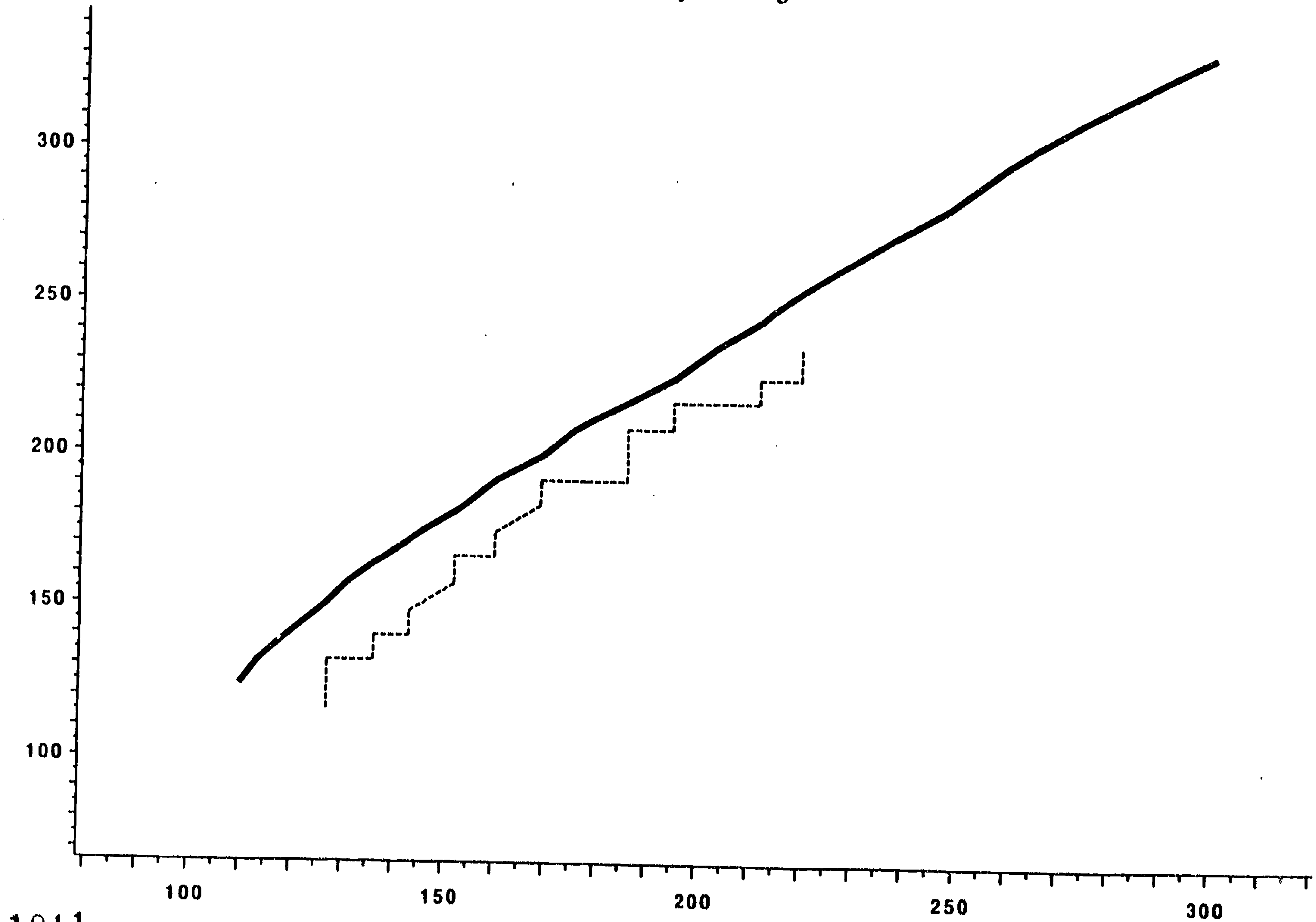
1009

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 62

English Language Spring K vs. Spanish Languagc Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



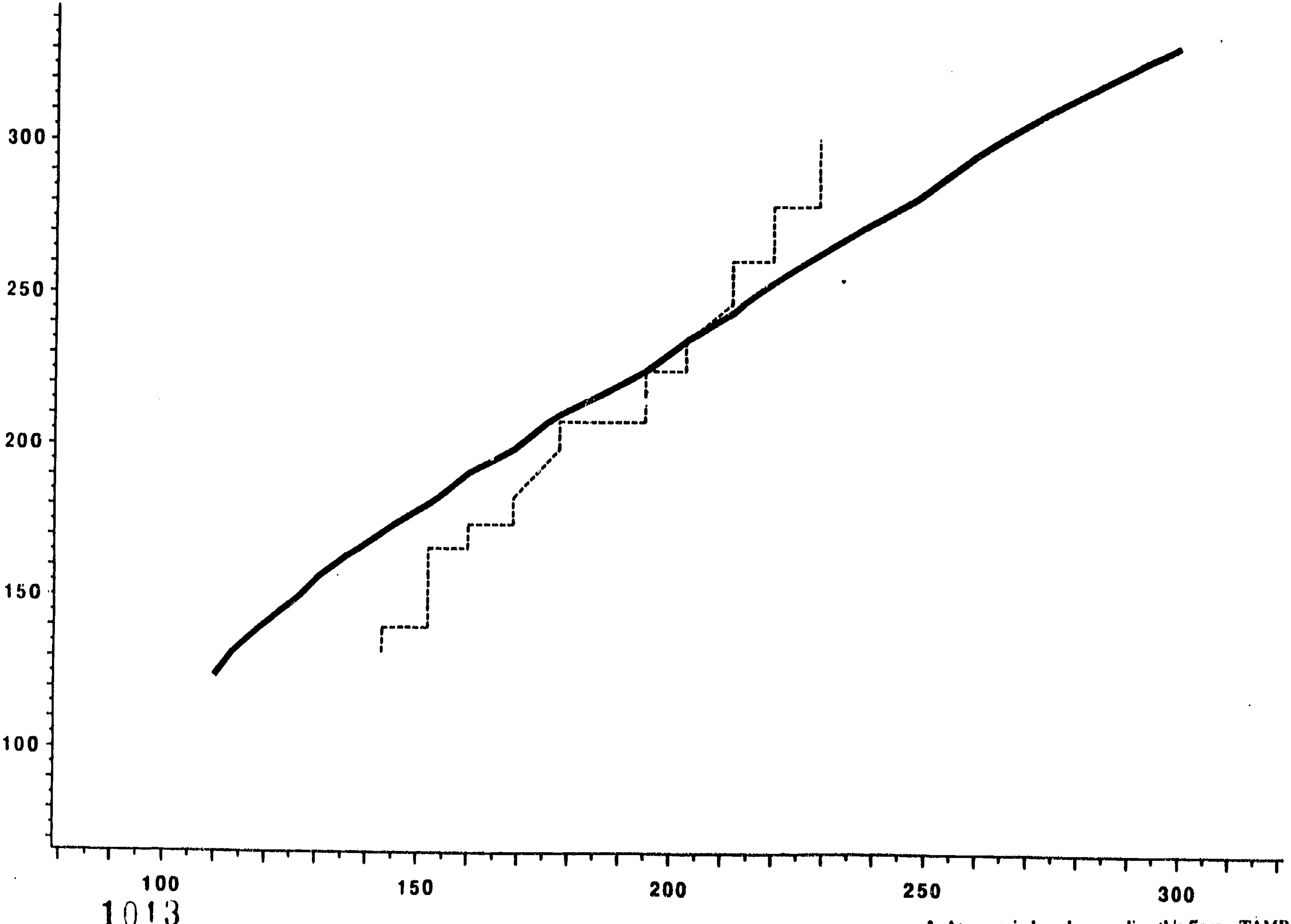
CURVE —— NORM - - - - - EE-C

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 63

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

2. Spring Kindergarten to Spring First Grade:

The TAMP analyses that follow differ from those presented earlier in that the time period has changed from fall kindergarten/spring kindergarten to spring kindergarten/spring first grade. The major shift is that the language used for measuring skills in kindergarten has changed from Spanish to English. The Spanish to English comparison made earlier was necessitated by the availability of only test data measured in Spanish and by the need to consider the hypothesis that students with higher initial primary language skills would perform better when tested in English than students with lower initial primary language skills. However, there is also a competing hypothesis that students exposed to more instruction in the second language will learn more than students taught mostly in their primary language. The change in language of test also allows us to examine the growth in achievement in English from kindergarten to first grade and to see how this growth pattern is similar or different from that describing the growth when measured in Spanish and English. Thus the comparisons in the following TAMP analyses cover English to English.

A.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

These analyses were effected to determine the comparability of mathematics skills relative to this norming population for immersion strategy, early-exit, and late-exit students at the end of kindergarten when tested in Spanish and at the end of first grade when tested in English.

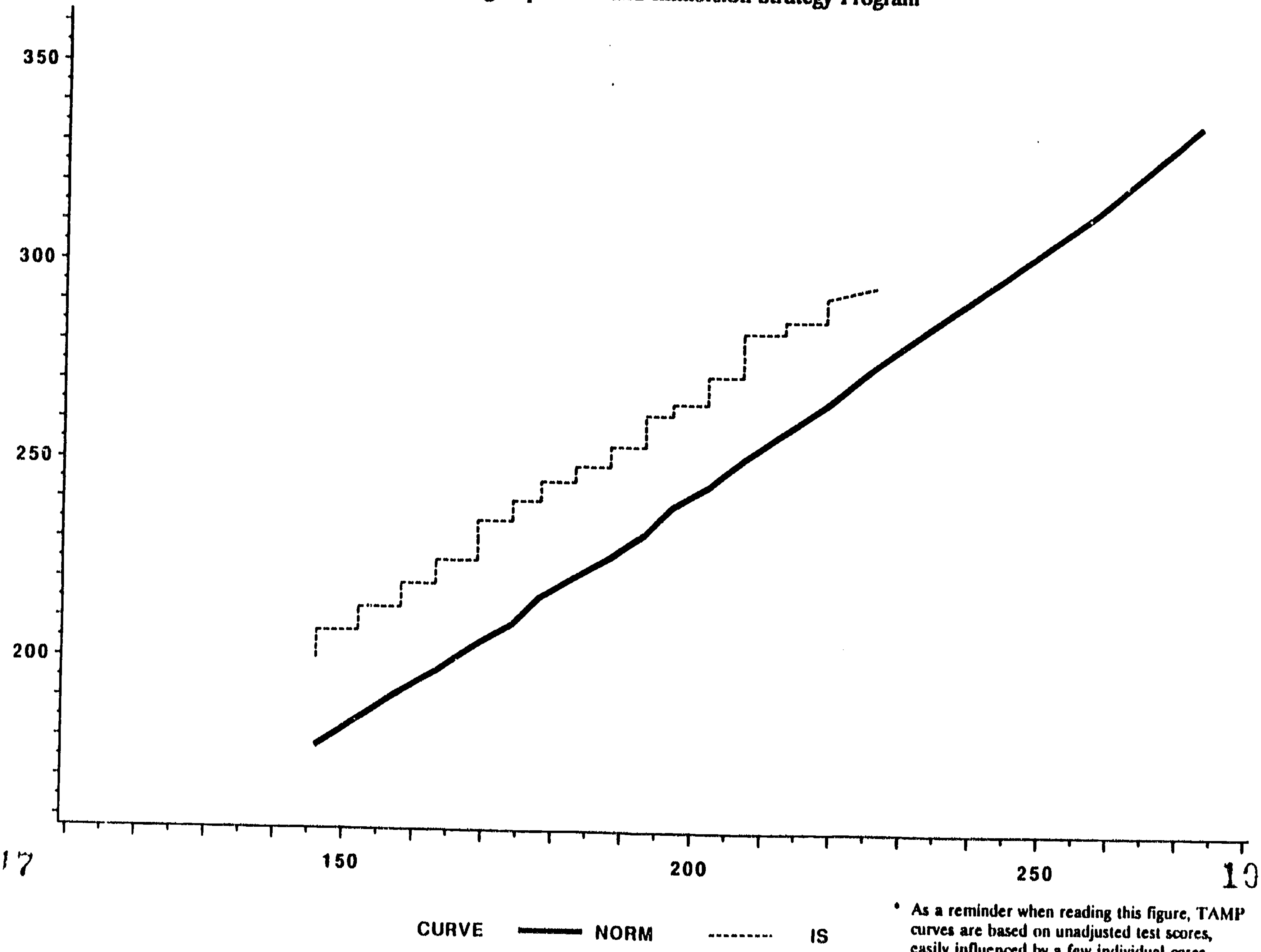
Figures 64, 65, and 66 suggest that although all three student groups in the TAMP curves ended their kindergarten year with lower scores relative to this norming population, the mathematics skills of each student group increased at a higher rate relative to this norming population when tested in English. This would suggest that if this growth rate were sustained over time these students would catch up to the average achievement of this norming population. However, they must be considered in the context of slower growth from fall kindergarten to spring kindergarten; i.e., they fell behind and now they are catching up. These growth rates are similar and consistent between sites within the immersion strategy program (see Figures 67 to 70). Among early-exit programs, students in sites EE-B and EE-C were growing faster than, and students in site EE-H as fast as, this norming population (see Figures 7. to 74).

It seems that each student group performed better than expected relative to this norming population, with immersion strategy students apparently showing growth at a higher rate than either early-exit or late-exit students. Both early-exit and late-exit students appear to have grown at the same rate relative to this norming population. This was predicted to occur as a result of the almost exclusive use of English in the immersion strategy program. That is, when mathematics skills are tested in English, those students who had the most instruction in English would perform the best. Also, their Spanish mathematics scores in spring would be relatively low. Contrary to expectations, the growth in achievement of early-exit students mirrored that of late-exit students. This may reflect the fact that some early-exit programs did not use as much English as predicted.

Figure 64

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



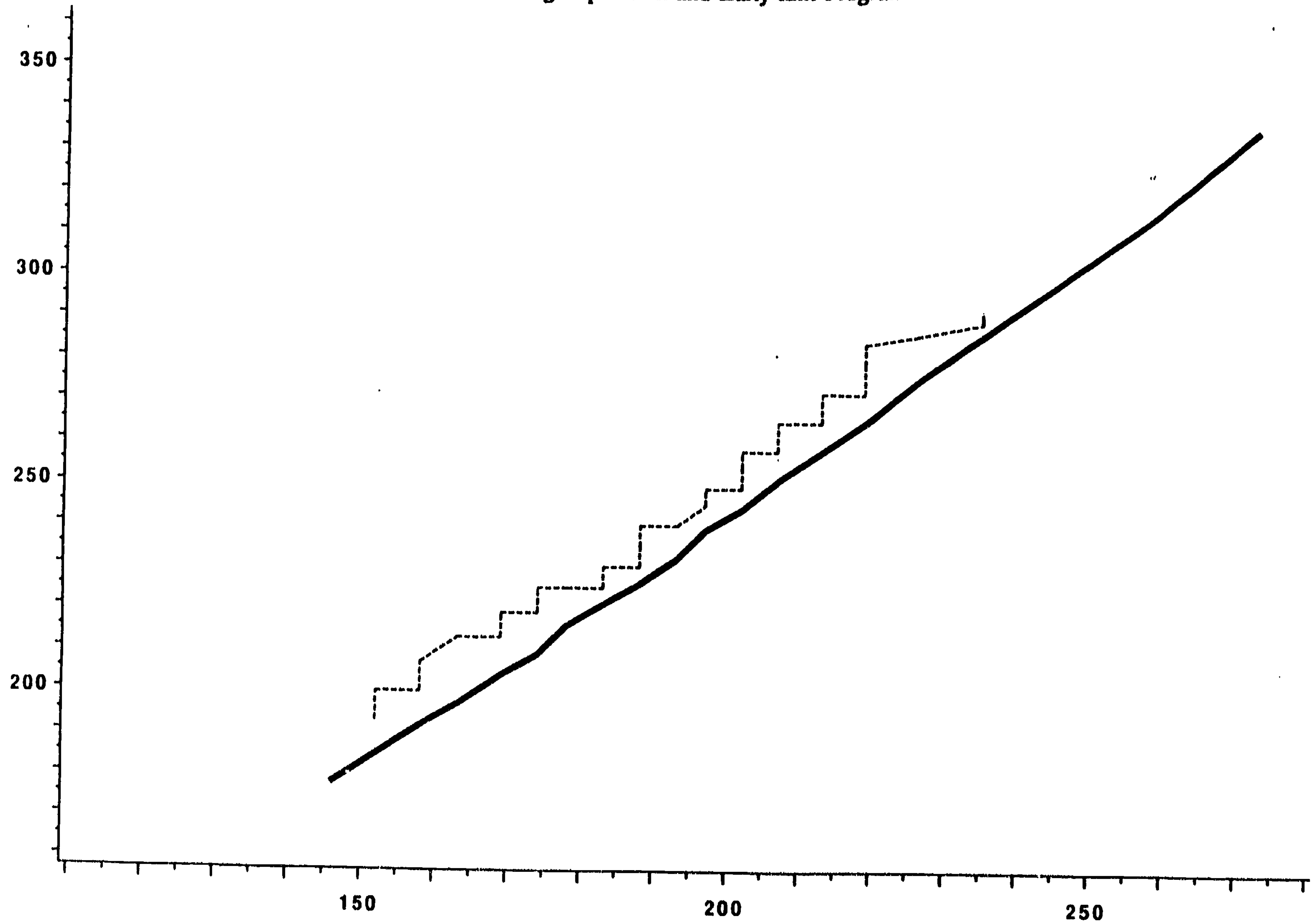
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

1017 1018

Figure 65

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



427

1019

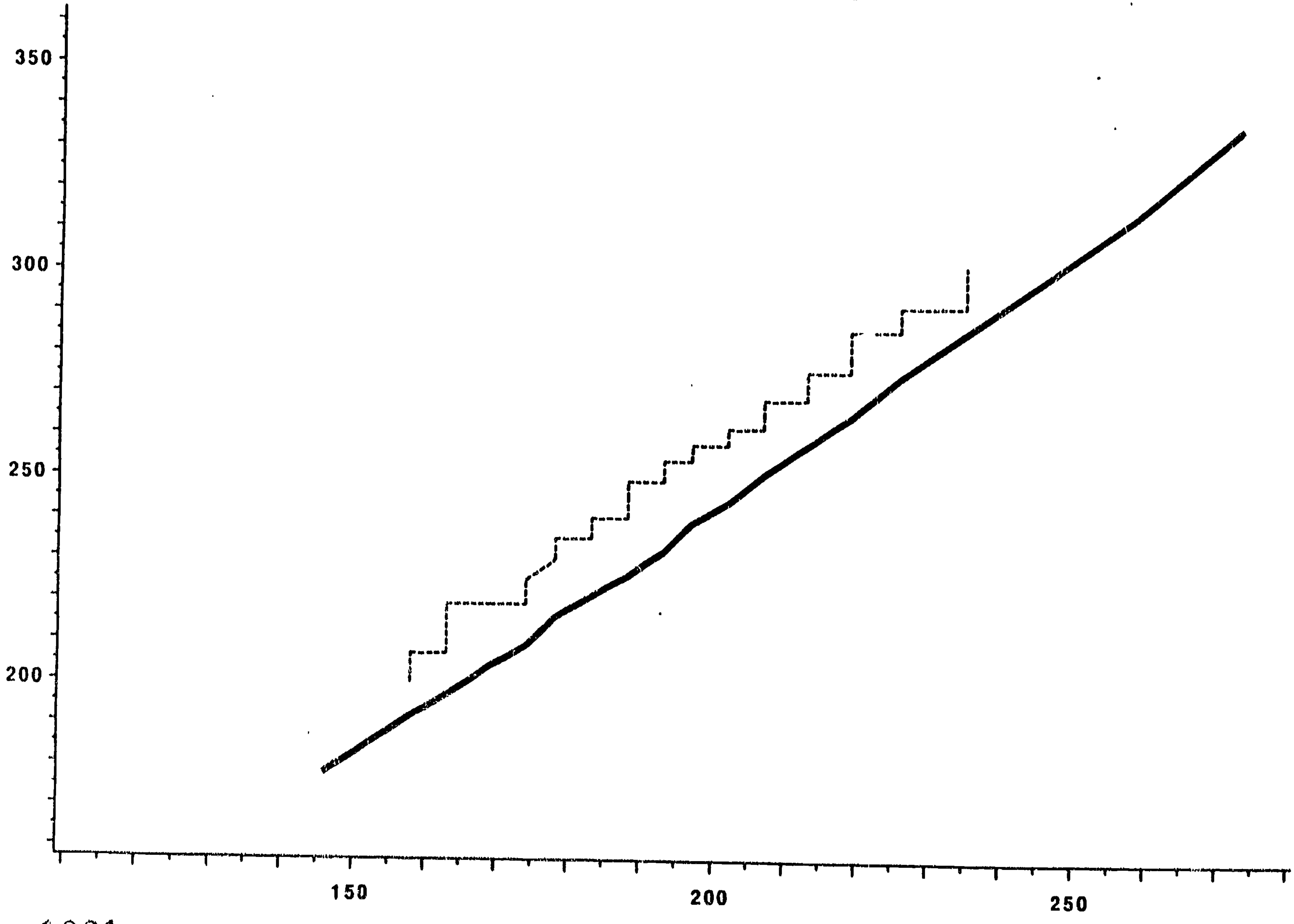
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 66

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



428

1021 1022

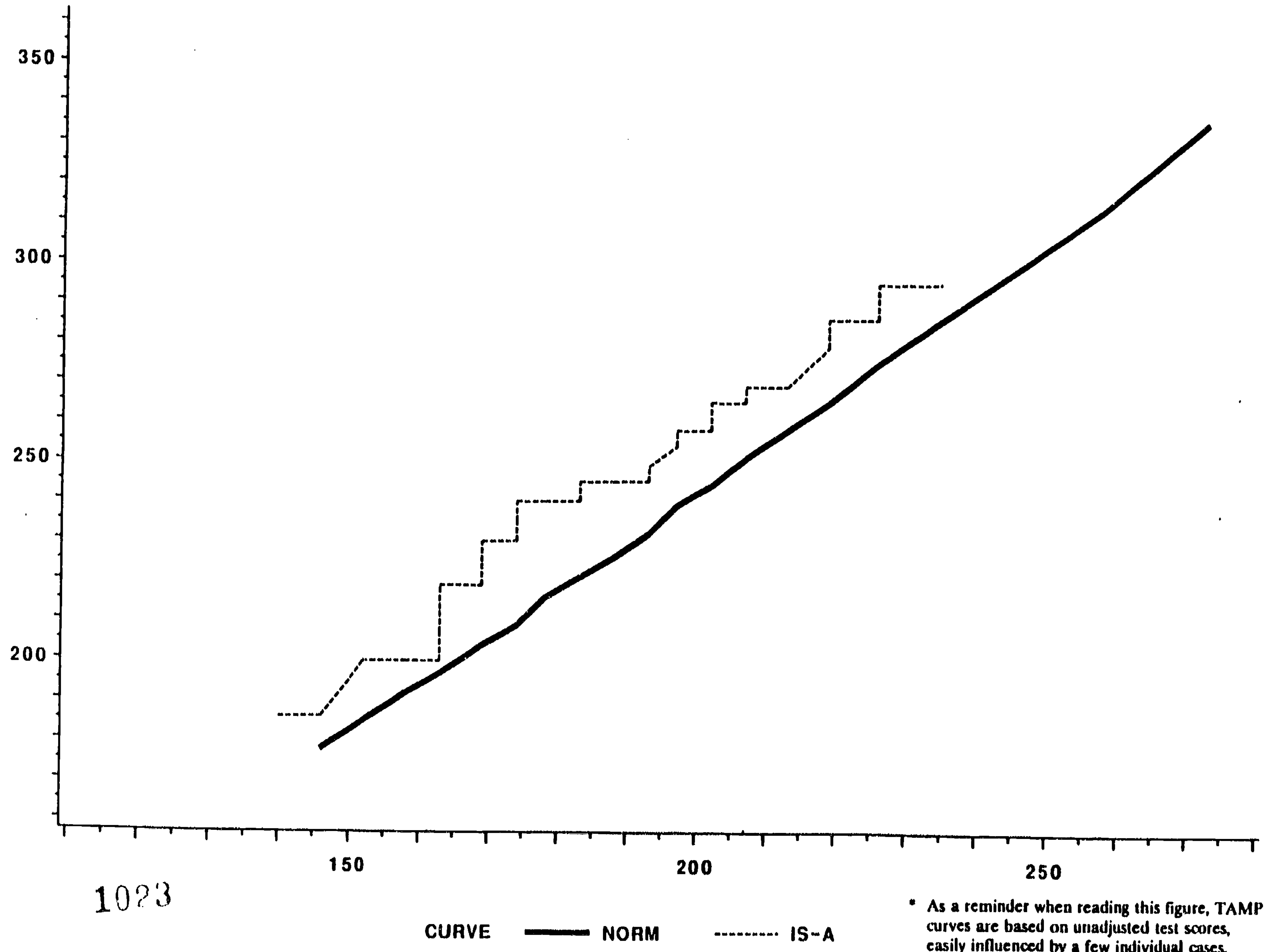
ERIC

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 67

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



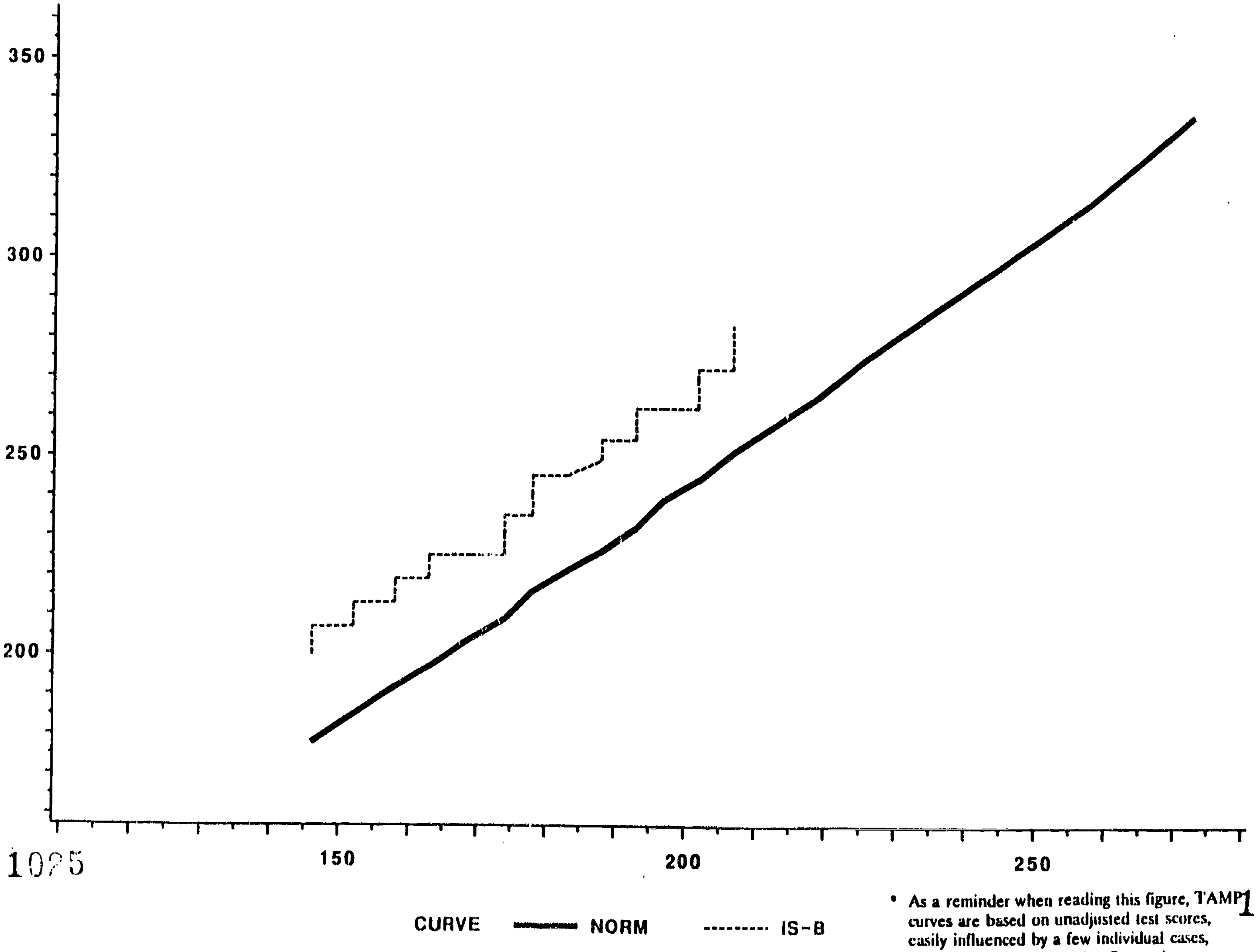
429

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 68

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



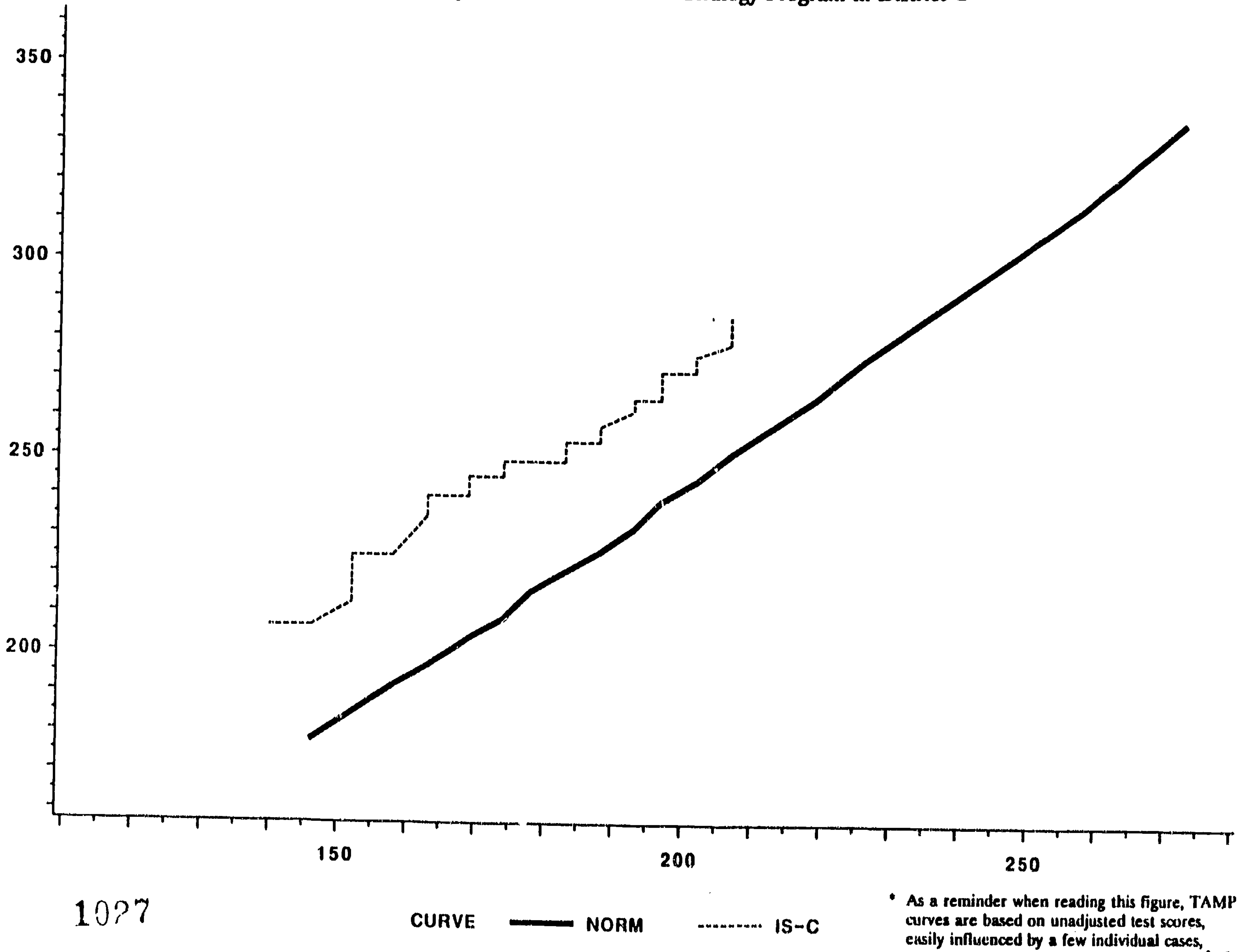
430

- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 69

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



431

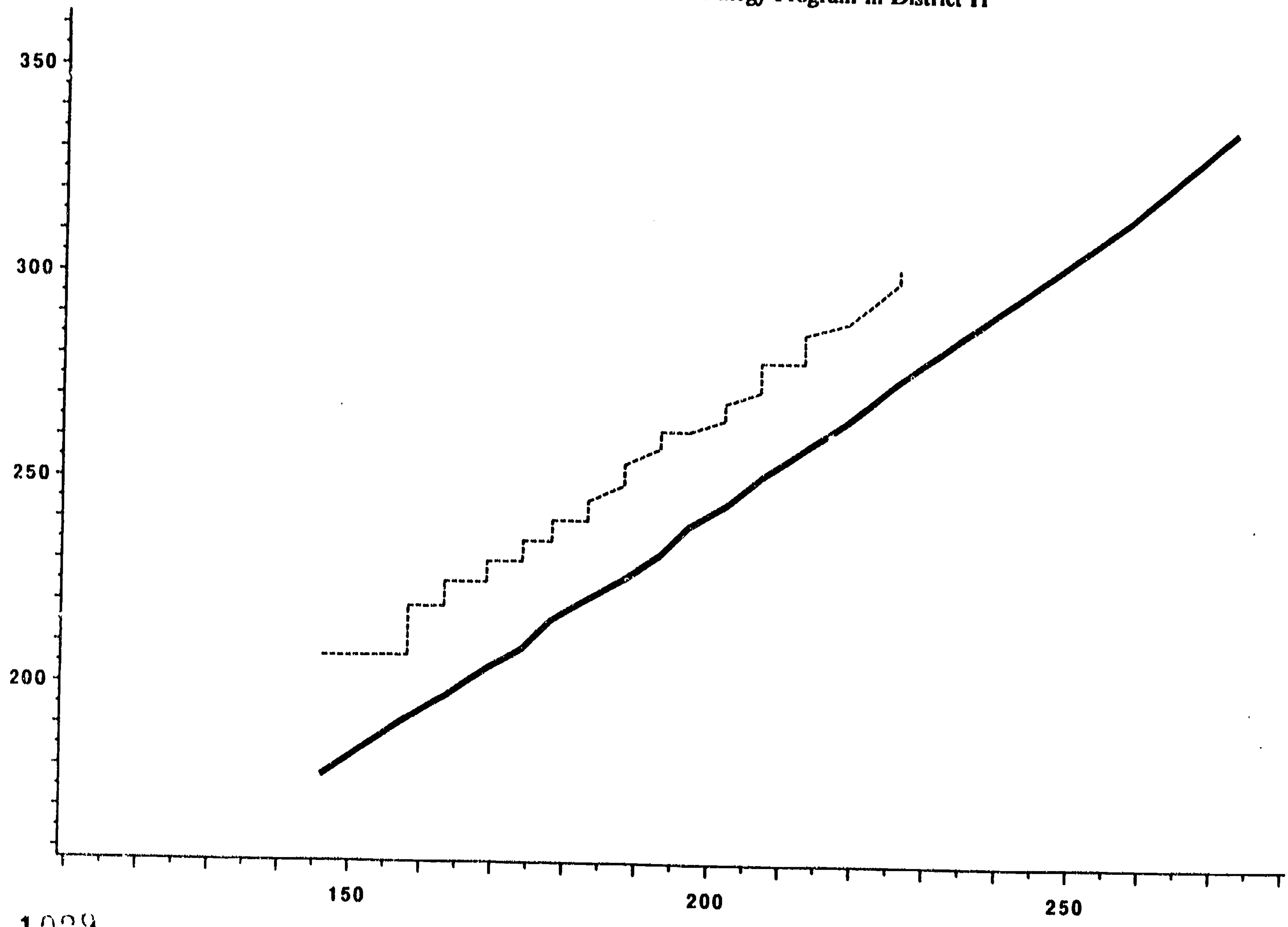
1027

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 70

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H

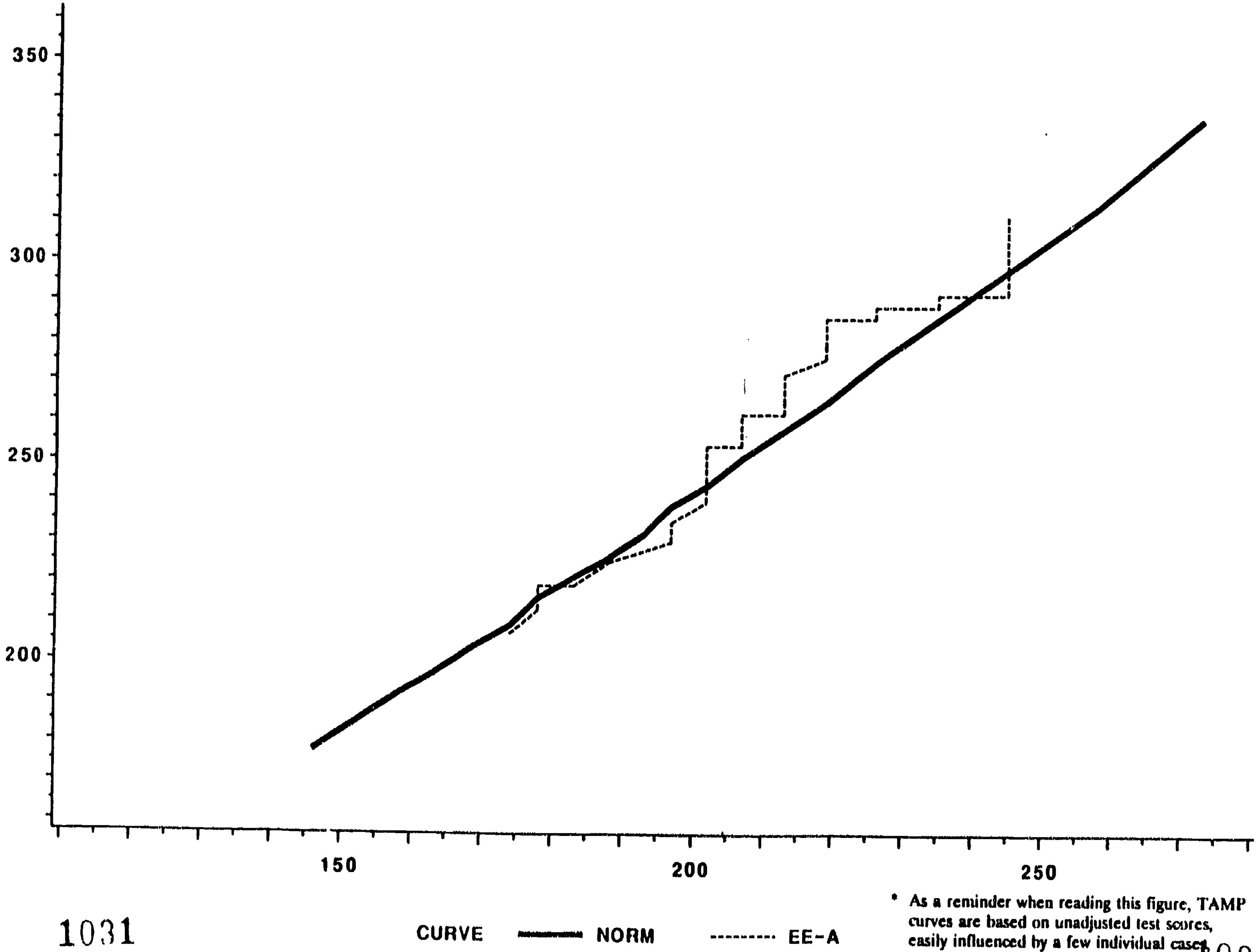


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 71

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



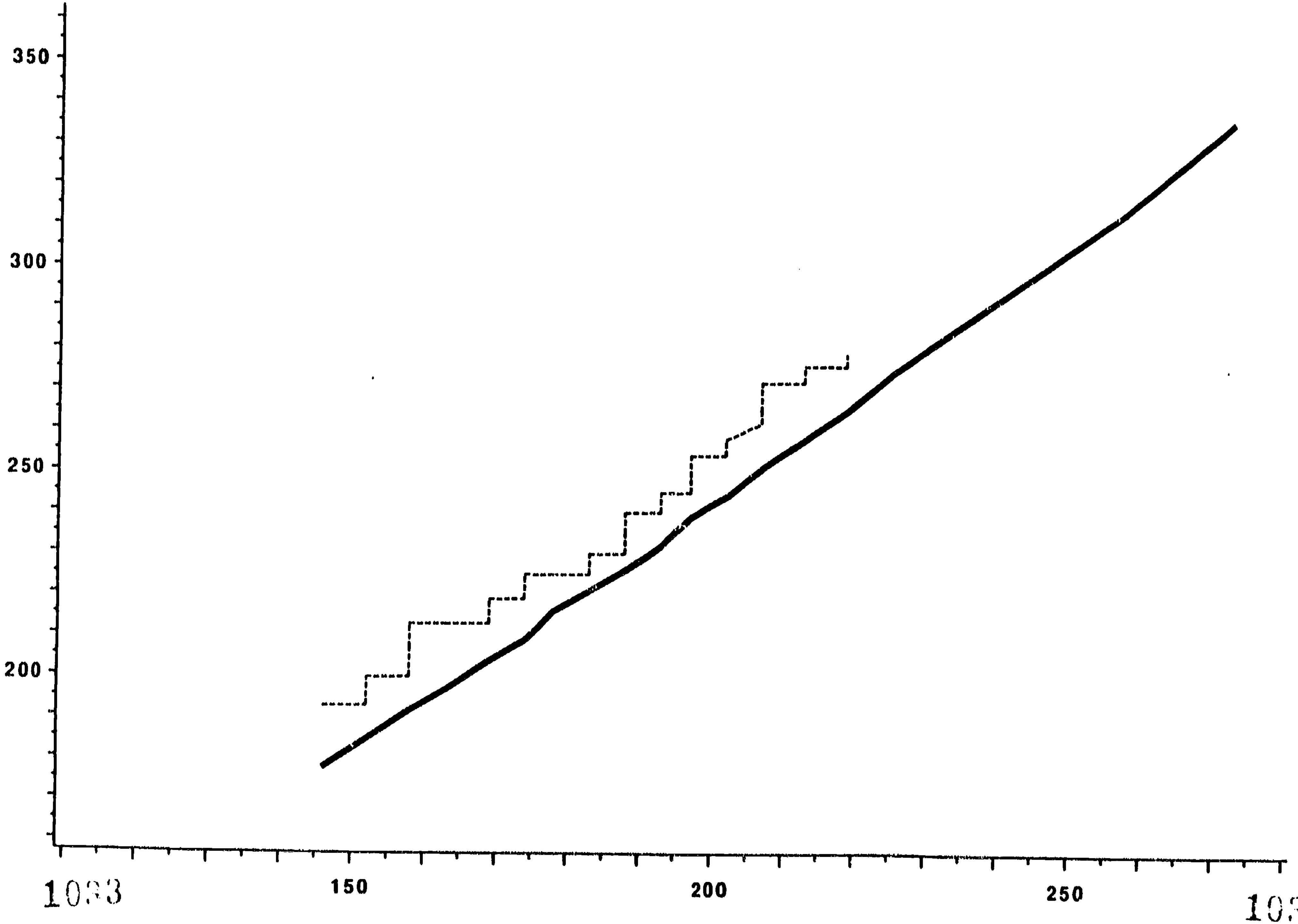
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases and are subject to sampling fluctuation.

ERIC

Figure 72

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B

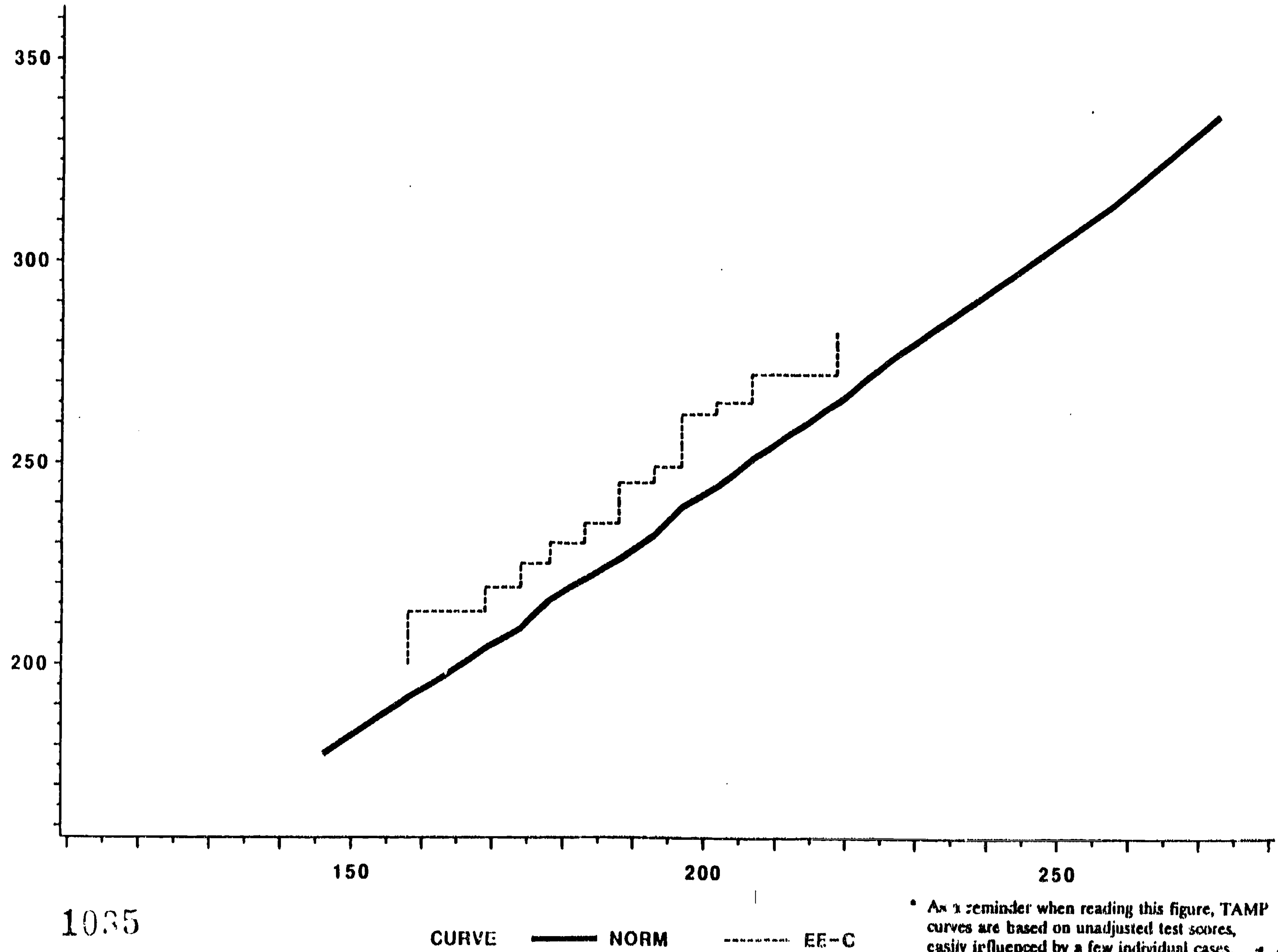


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 73

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



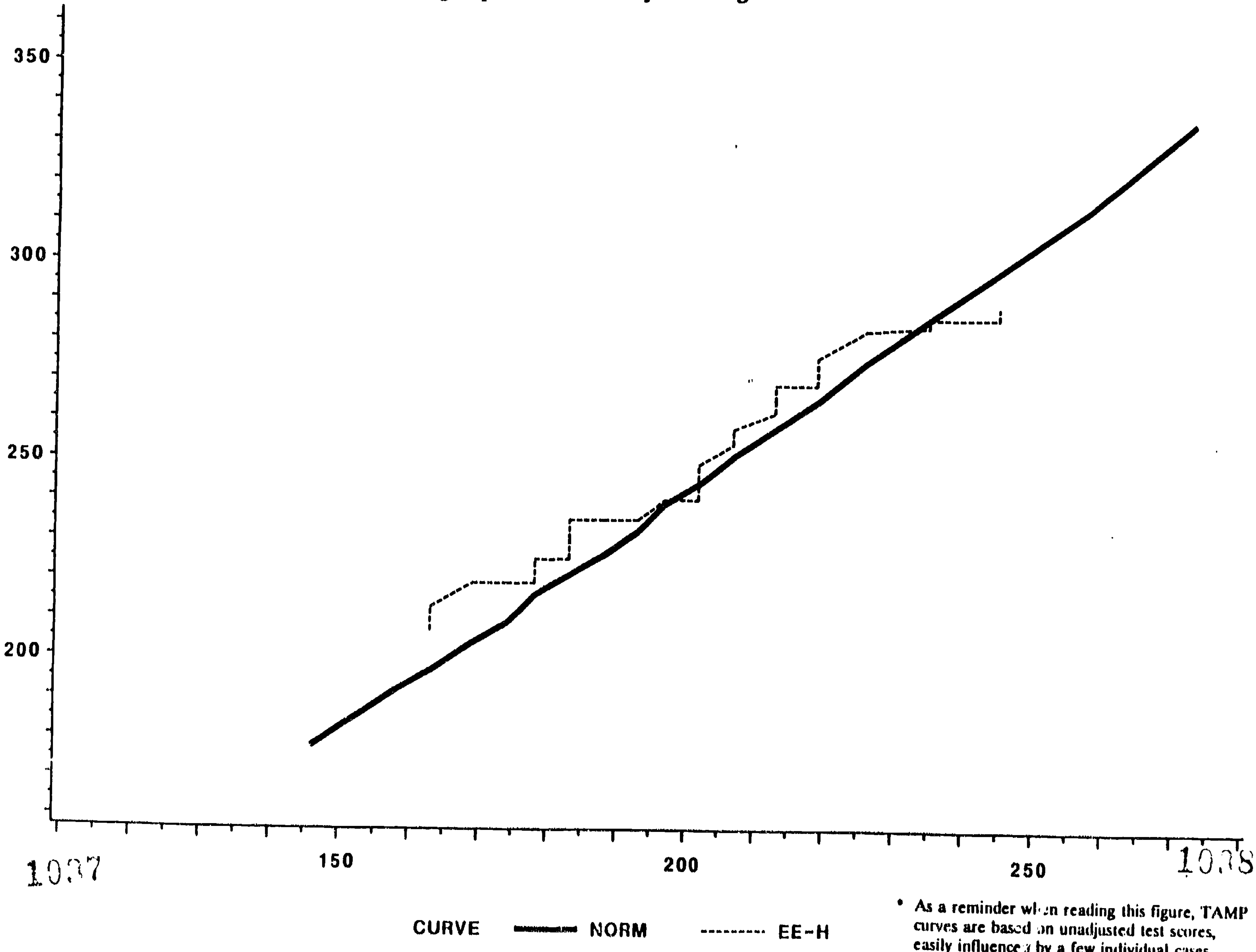
435

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 74

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

1037 1038

B.	Grade Span:	Kindergarten	to	First Grade
	Test Date:	Spring	to	Spring
	Language:	English	to	English
	Content:	Mathematics	to	Mathematics

As above, these analyses were done to understand the growth rates of immersion strategy, early-exit, and late-exit students relative to this norming population. Students in all three student groups appear to have higher growth rates in mathematics when tested in English than this norming population (see Figures 75, 76, and 77). Immersion strategy and late-exit students seemed to increase their mathematics skills at about the same rate, with early-exit students not quite as fast. Moreover, these results were consistent across languages (i.e., regardless of whether their spring mathematics skills were assessed in English or Spanish).

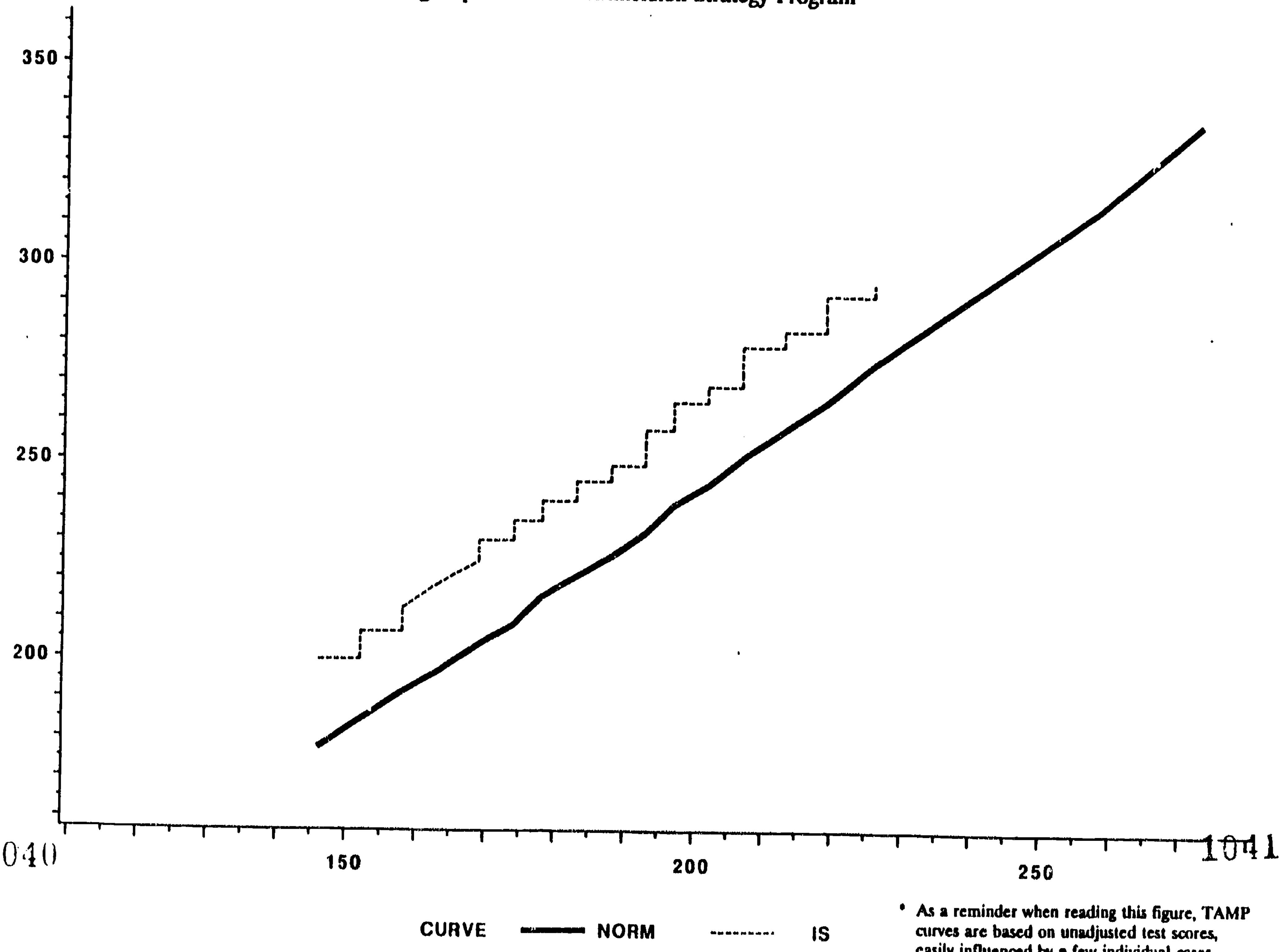
When the individual TAMP curves for each site were examined, some within-program variation was noted. With the exception of immersion strategy site C, immersion strategy students seem to have grown at a slightly higher rate than this norming population (see Figures 78 to 81); site IS-C appeared to grow at a noticeably faster rate than this norming population. There was a great deal of variation within the early-exit program, as students in site EE-C grew at a faster rate and those in site EE-H grew more slowly than in the other early-exit sites (see Figures 82 to 85). Moreover, students in site EE-A began with higher initial mathematics skills when tested in English than in either site EE-B or EE-C. This may reflect that EE-B and EE-C students received the most instruction in Spanish in the first grade in the early-exit program. Nonetheless, all early-exit sites except EE-H seemed to grow faster than this norming population.

ERIC

Figure 75

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

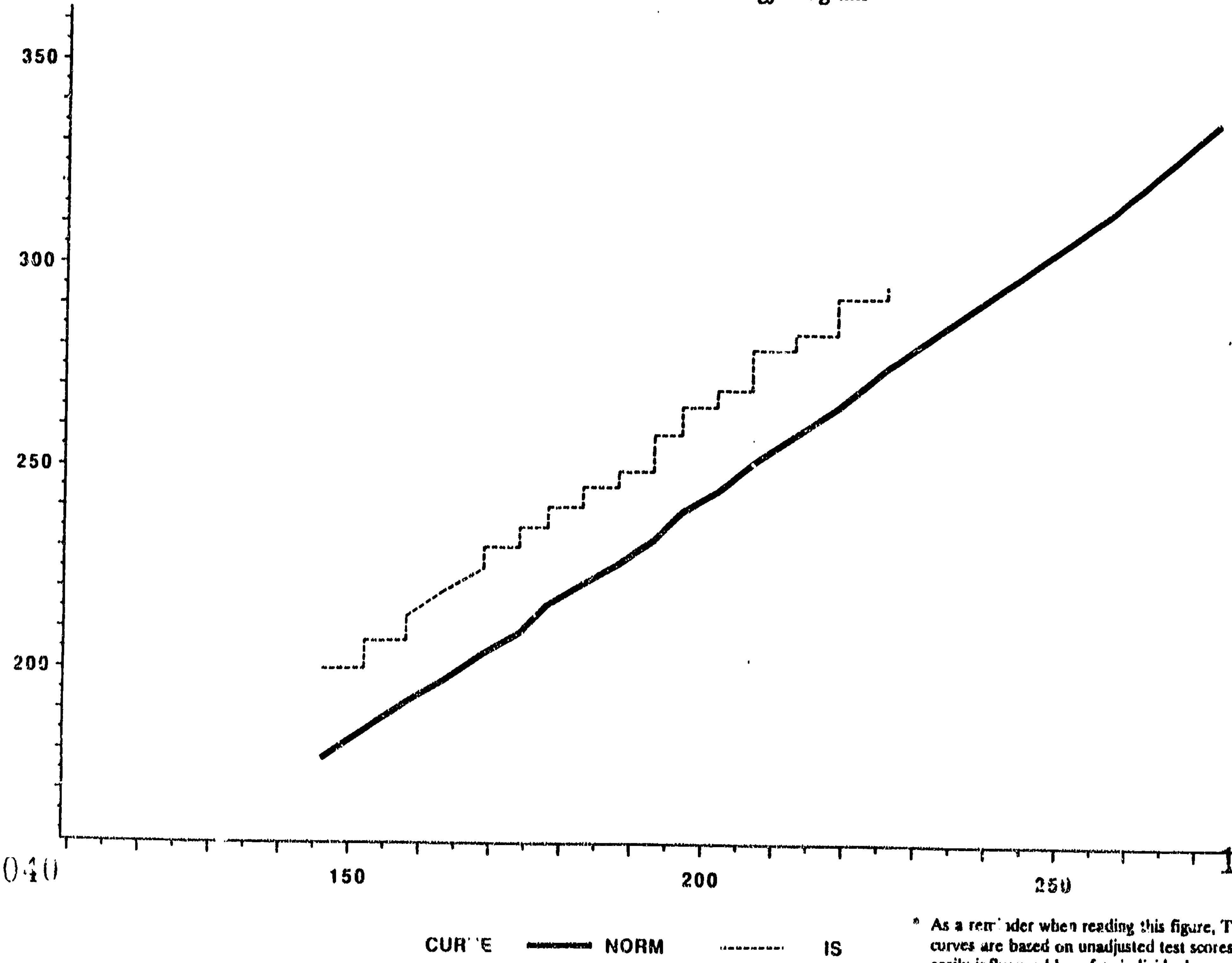


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 75

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



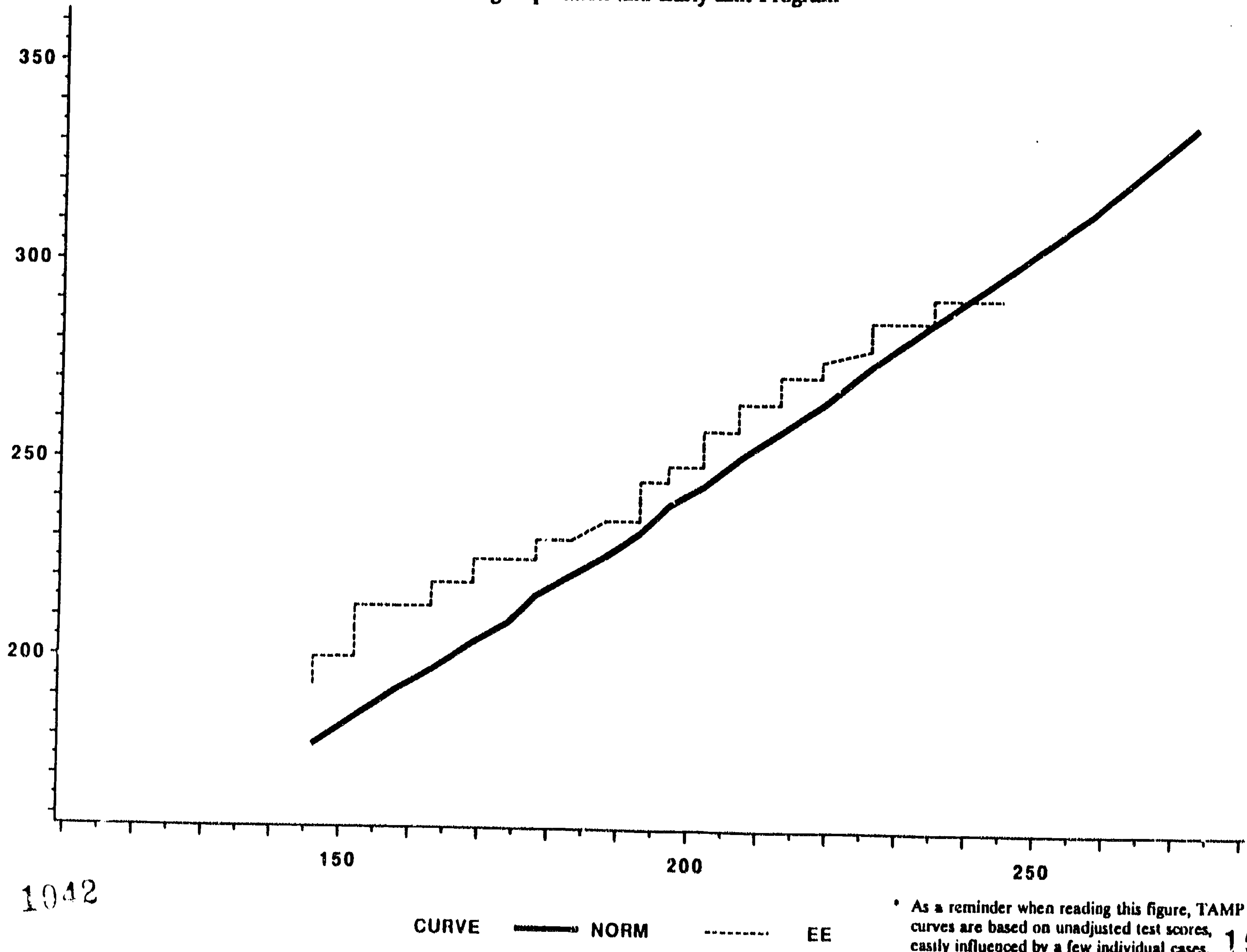
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 76

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

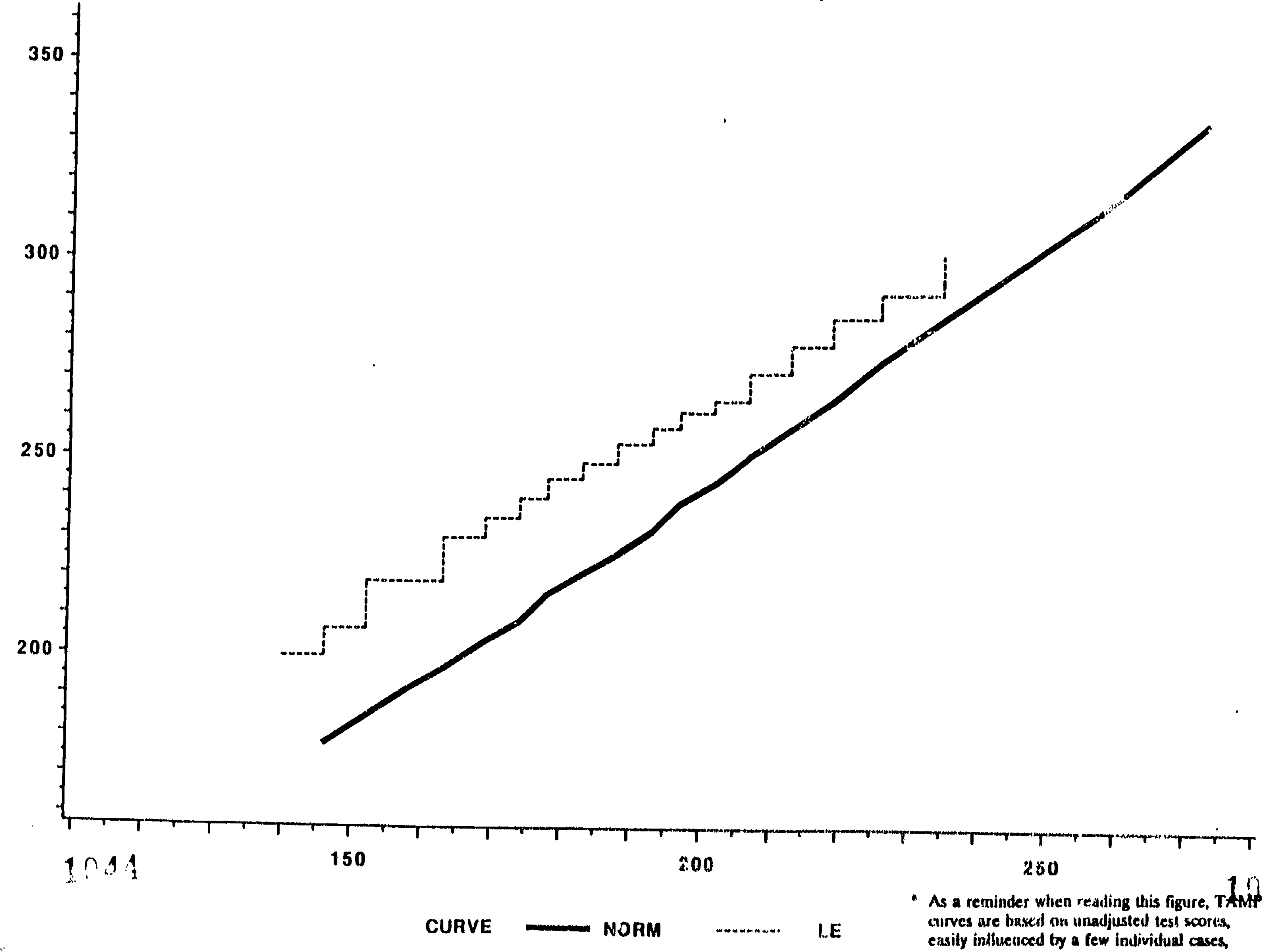


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 77

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program

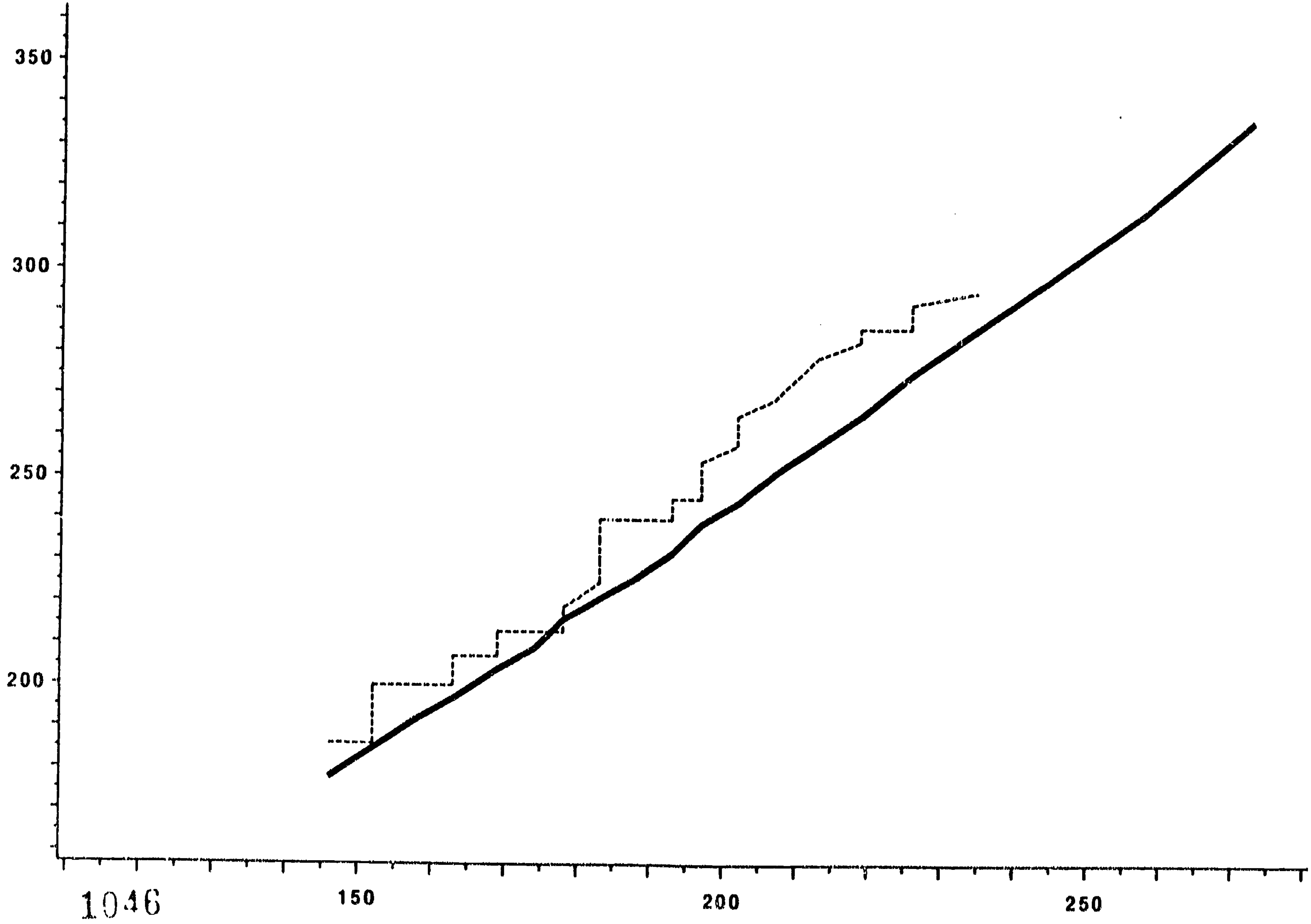


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 78

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



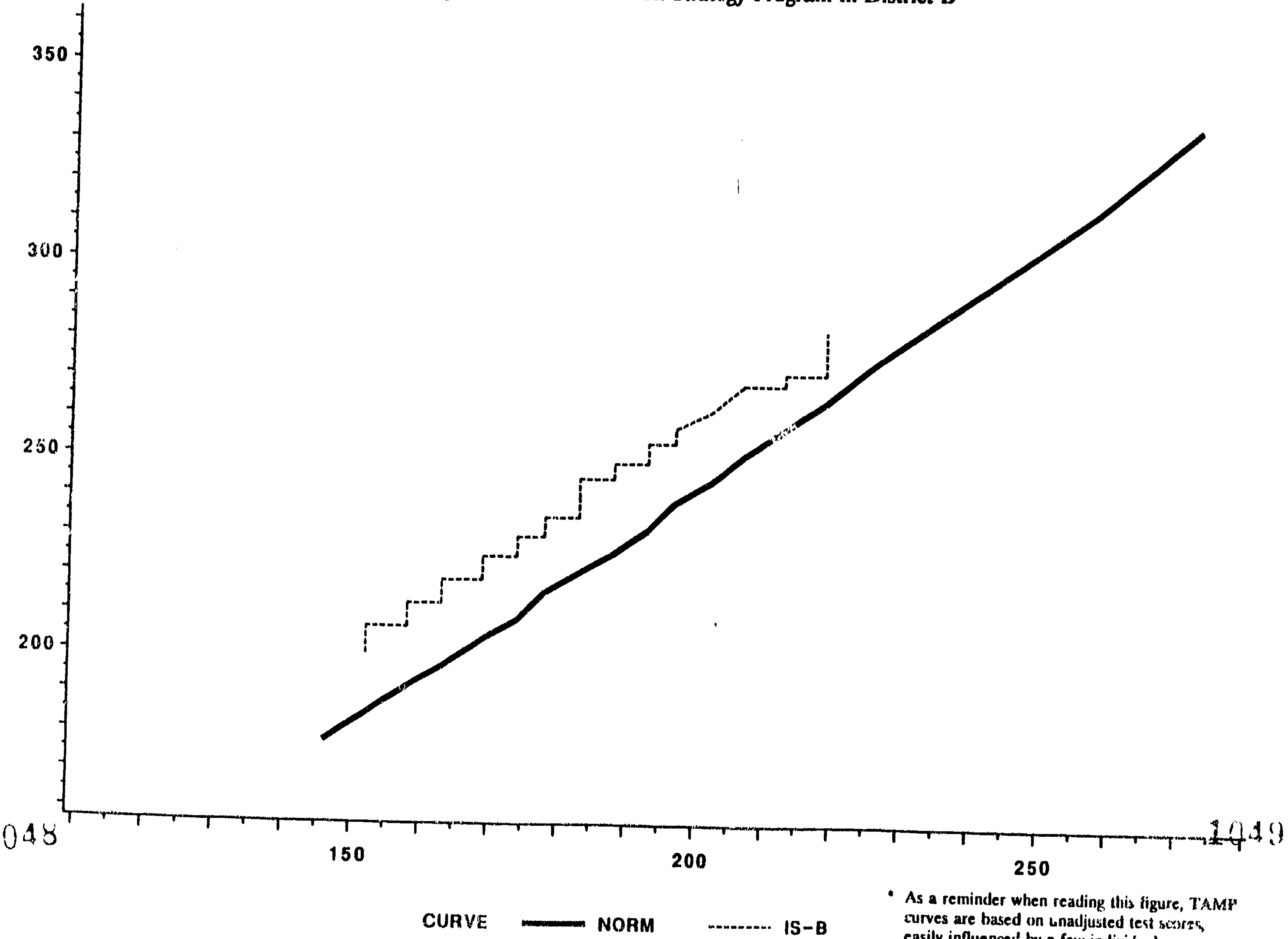
441

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 79

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



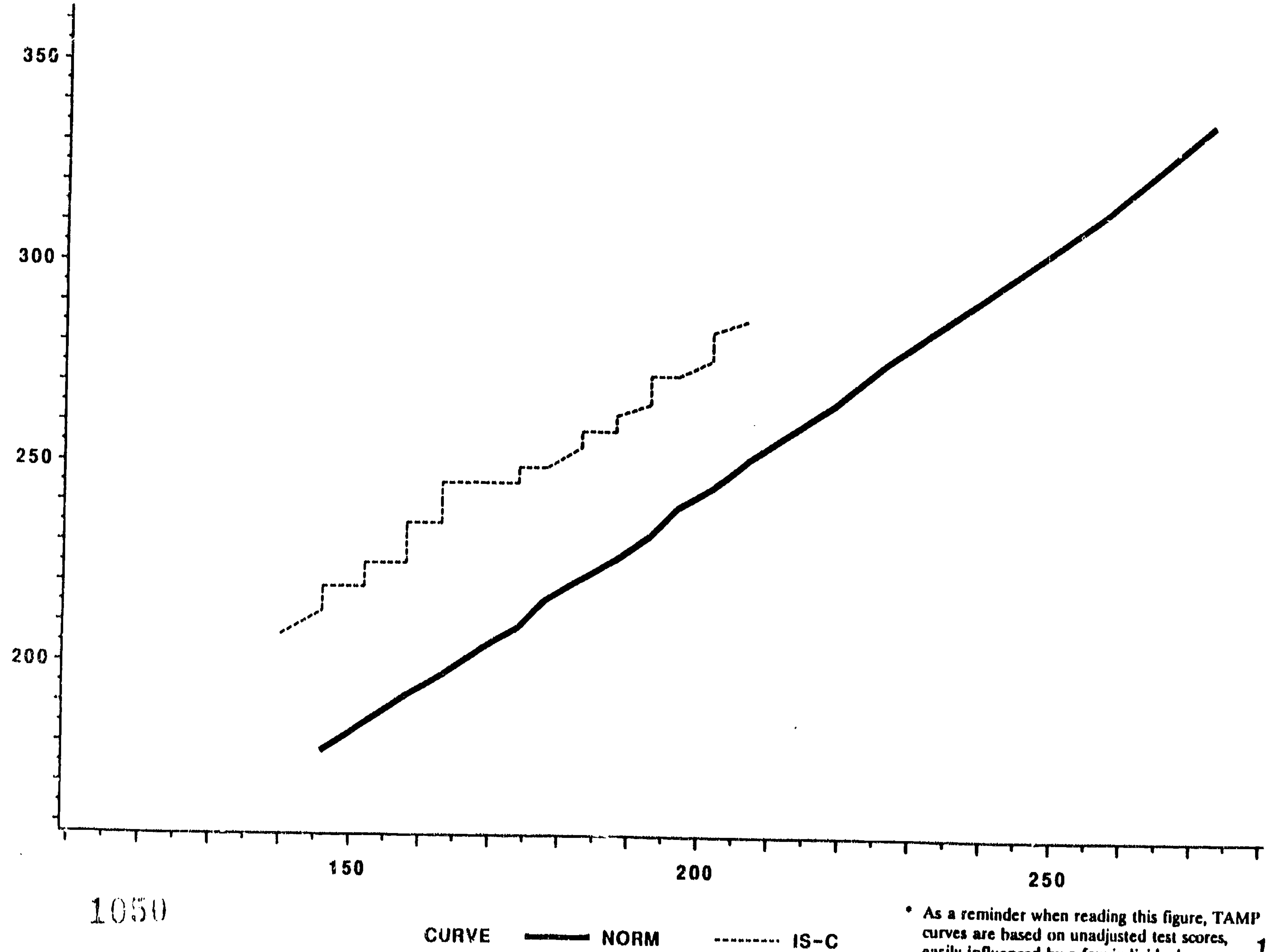
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

1048 1049

Figure 80

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



443

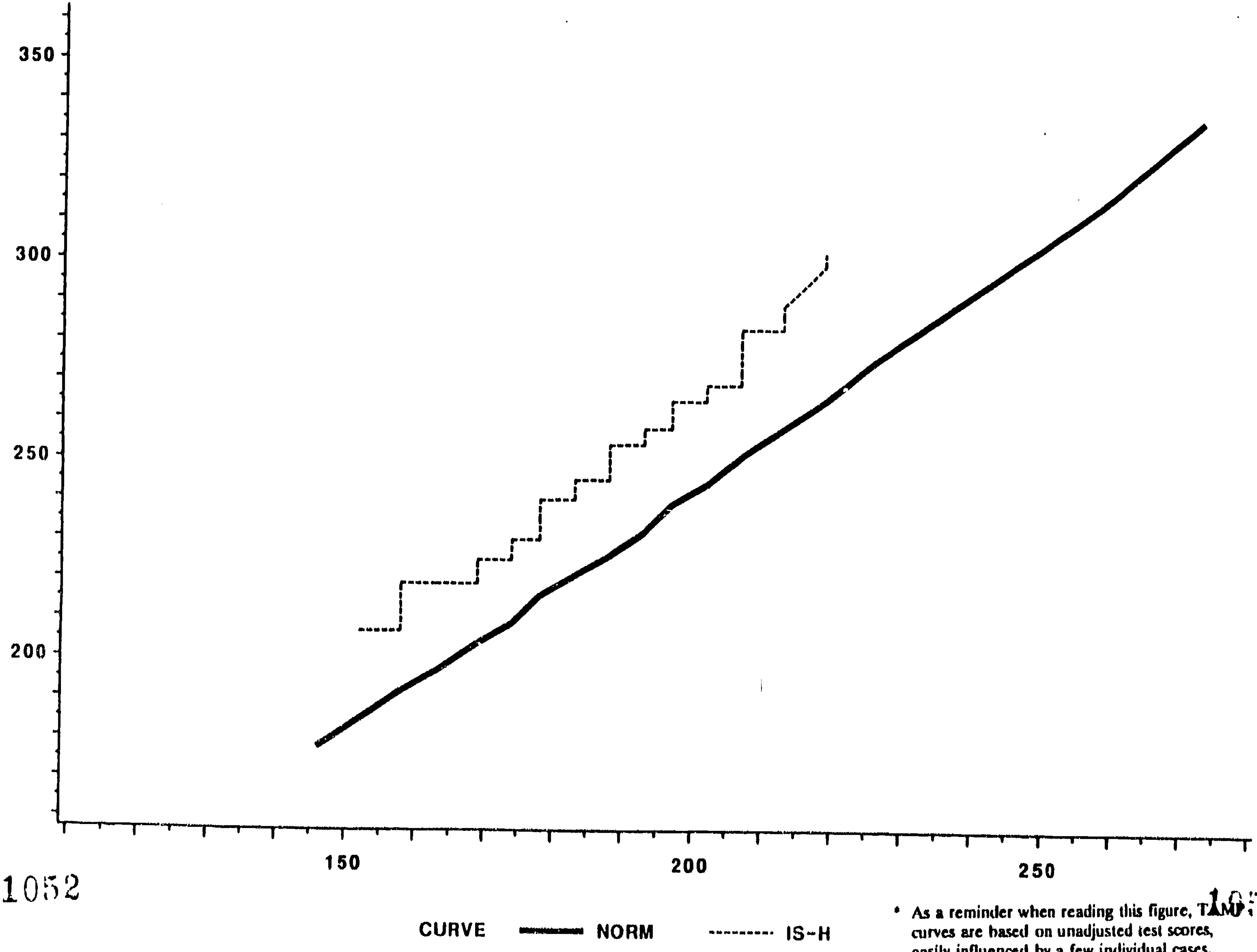
1050

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 81

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H



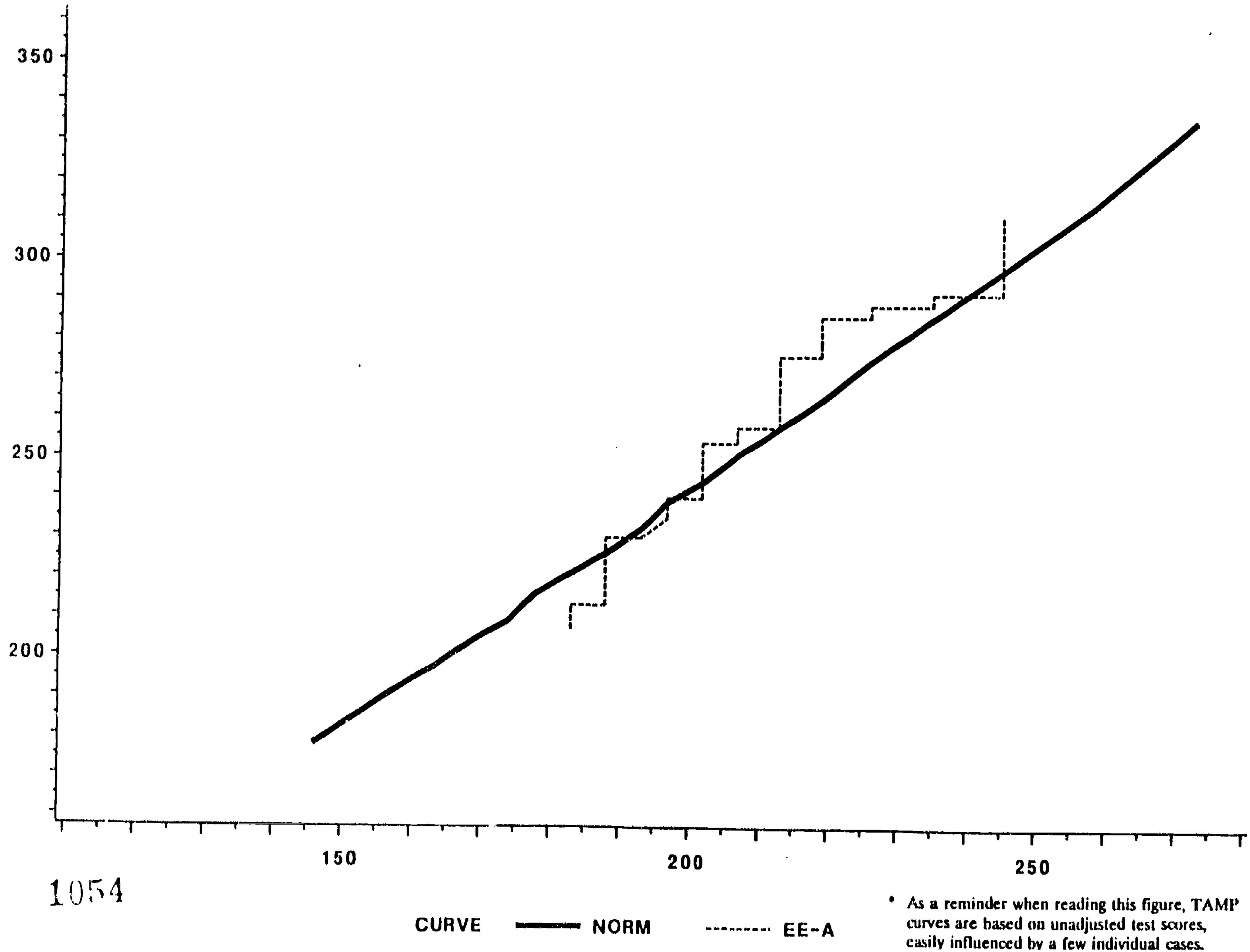
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 82

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



445

1054

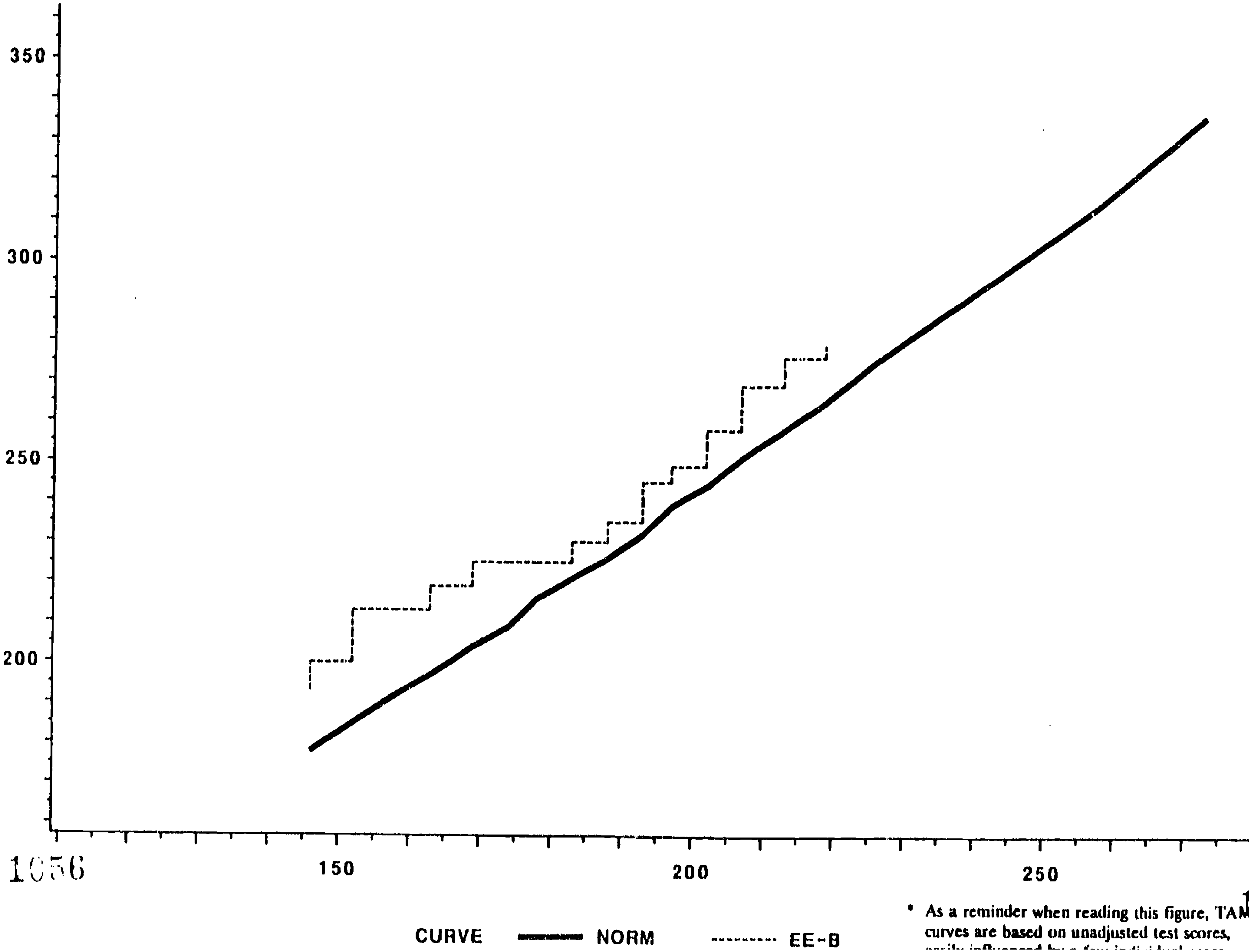
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

Figure 83

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



446

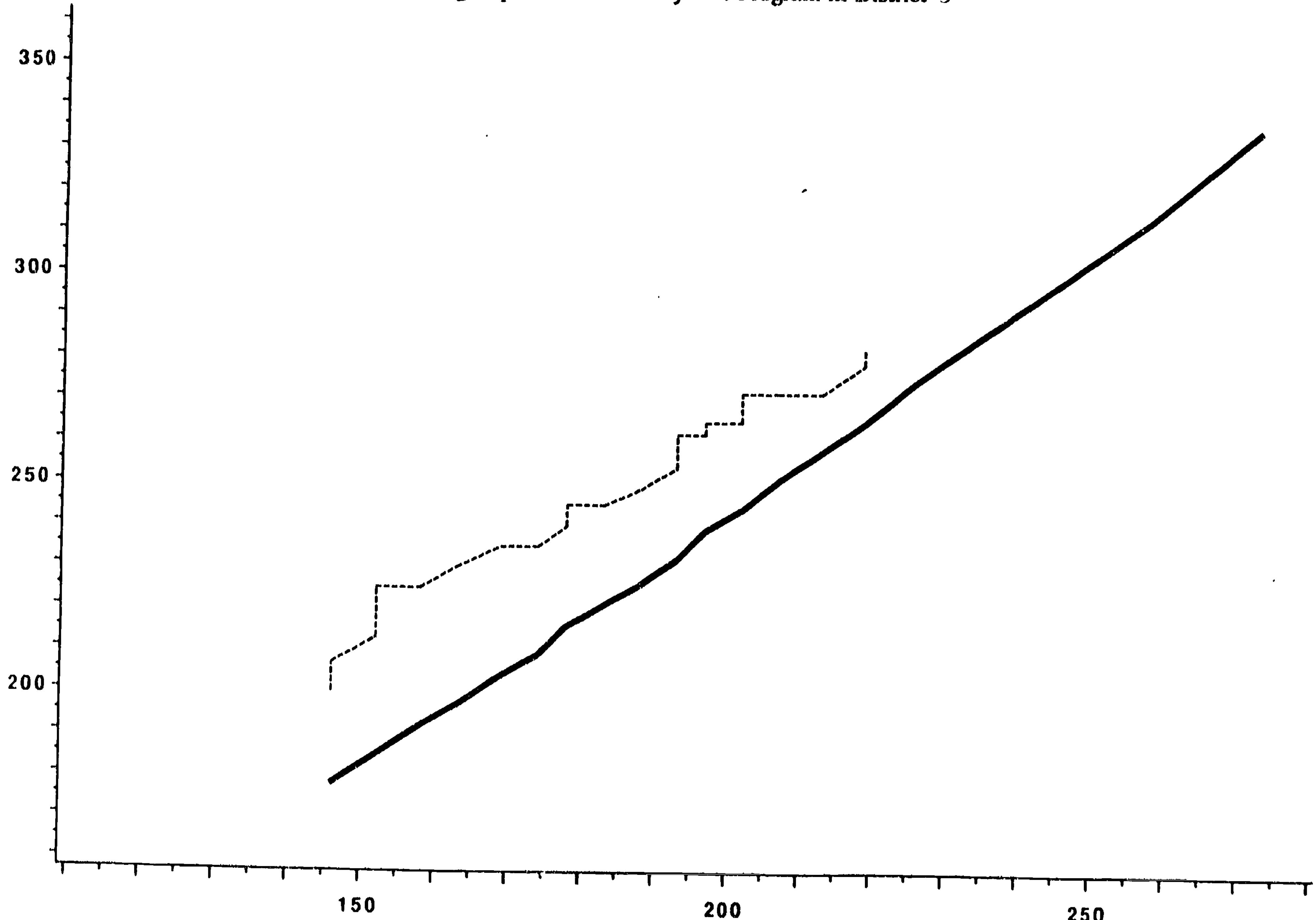
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores,

ERIC

Figure 84

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



447

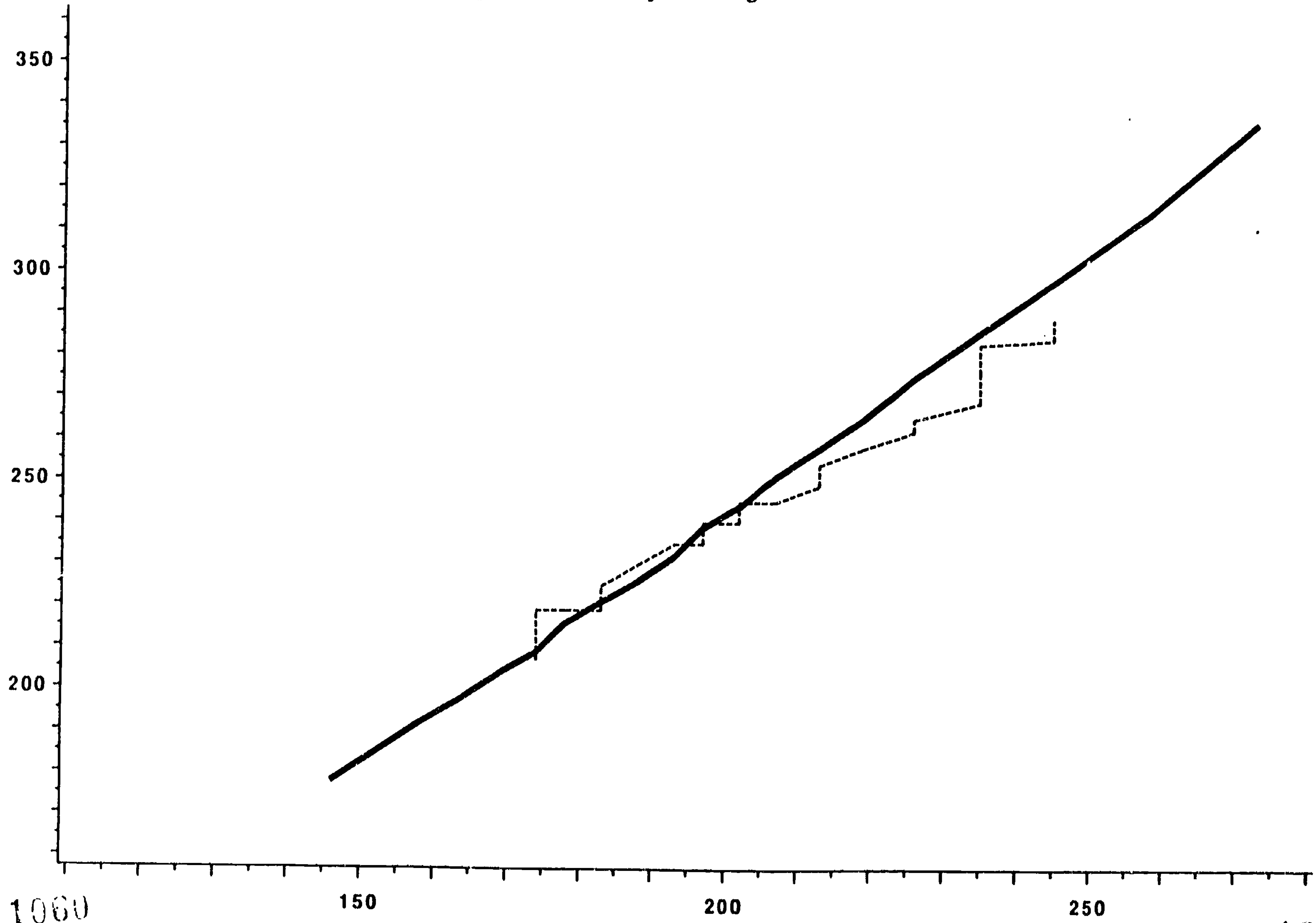
1058

• As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 85

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMI Curves: Norming Population and Early-Exit Program in District H



CURVE —— NORM - - - - EE-H

* As a reminder when reading this figure, TAMI curves are based on unadjusted test scores, easily influenced by a few individual cases,

ERIC

C.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Language**	**to**	**Language**

These analyses helped determine how the spring Spanish language skills assessed in kindergarten and the spring English language skills tested in first grade related to this norming population. At the end of kindergarten it seems that all three groups had slightly lower language skills relative to this norming population, but were more or less comparable to one another vis-a-vis this norming population (see Figures 86, 87, and 88). Immersion strategy students seemed to learn at a faster rate than this norming population, suggesting that they were catching up after exhibiting lower Spanish skills at spring kindergarten. Early-exit and late-exit students appeared to grow at the same rate as this norming population. As with mathematics, immersion strategy students appear to have grown at a faster rate than either early-exit or late-exit students relative to this norming population in English language skills.

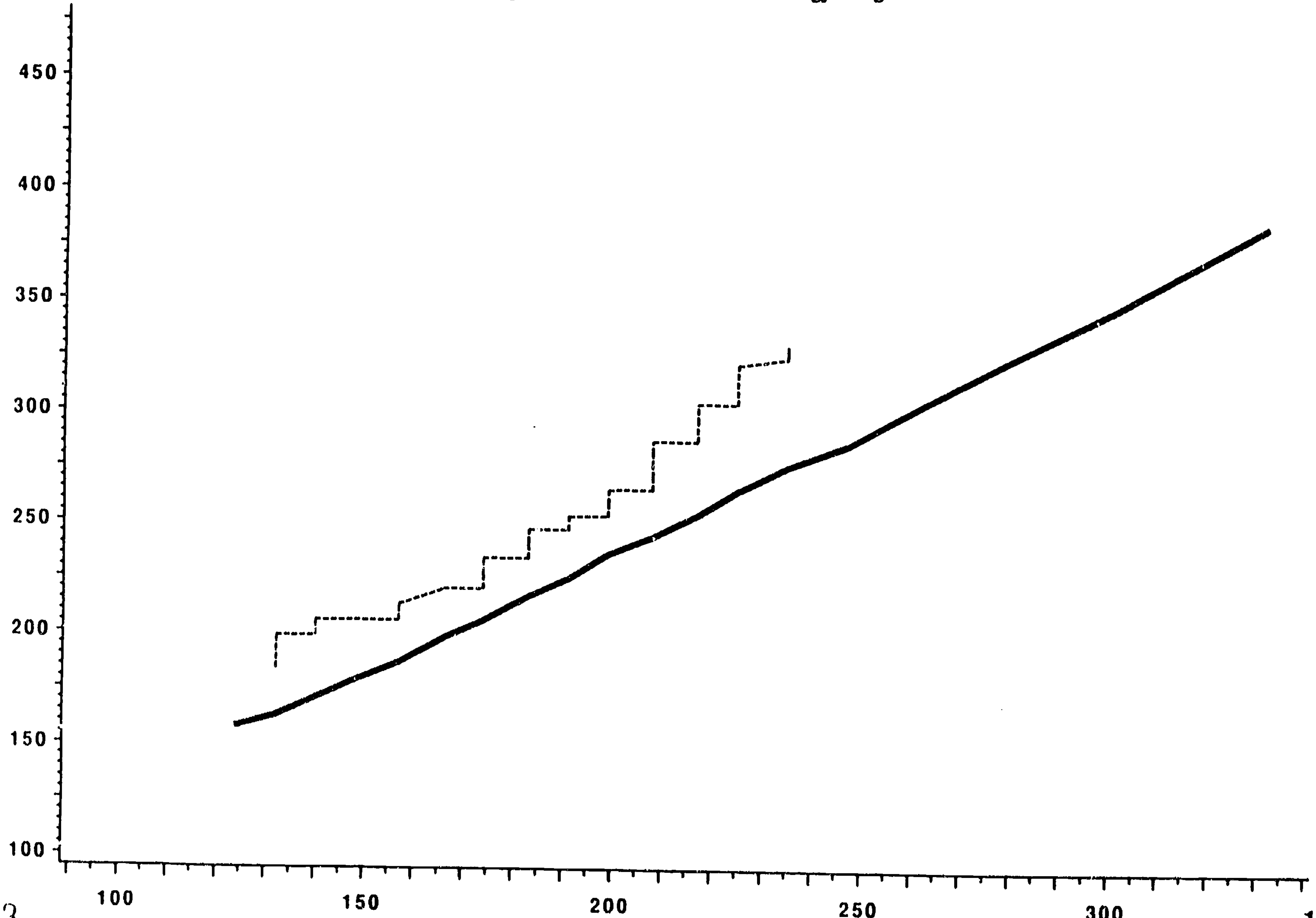
It also appears that the upward jog in growth in English skills exhibited by immersion strategy students noted in Chapter III occurred during their first grade year. Recall from earlier in this chapter that their growth in English language skills from fall kindergarten to spring kindergarten was slower (i.e., losing ground) than this norming population. In contrast, their growth in English language skills from spring kindergarten to spring first grade was greater than that of this norming population, indicating that they were catching up.

Finally, TAMP curves for each site by program did not reveal any material variation among the immersion strategy sites (see Figures 89 to 92). As before, there was slight variation between the four early-exit sites. Students in site EE-A with the highest initial skills seemed to learn at a much faster rate than this norming population, and site EE-H students as a group grew slightly slower than this norming population and students in the other early-exit sites (see Figures 93 to 96).

Figure 86

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

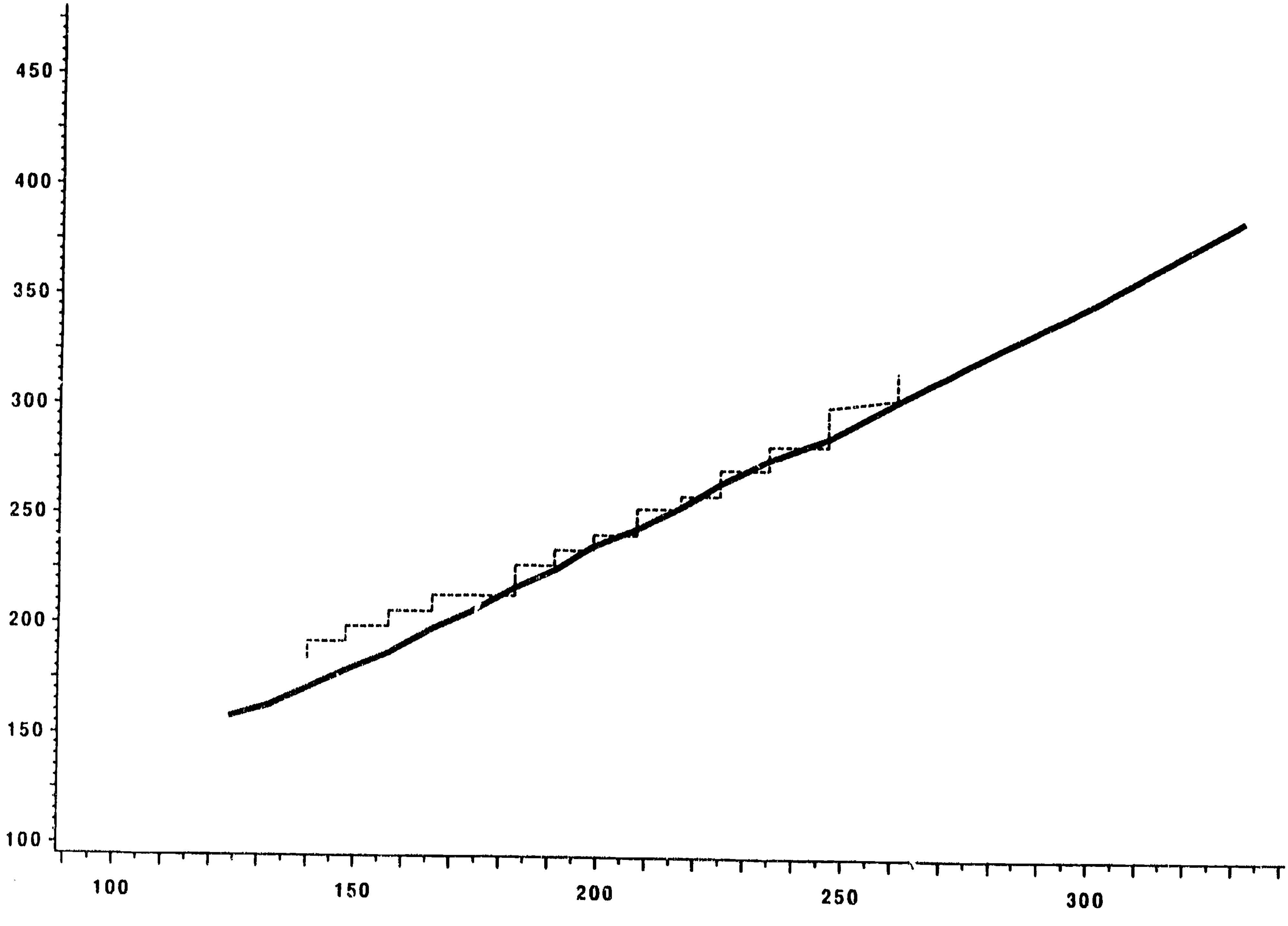


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 87

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



451

1065

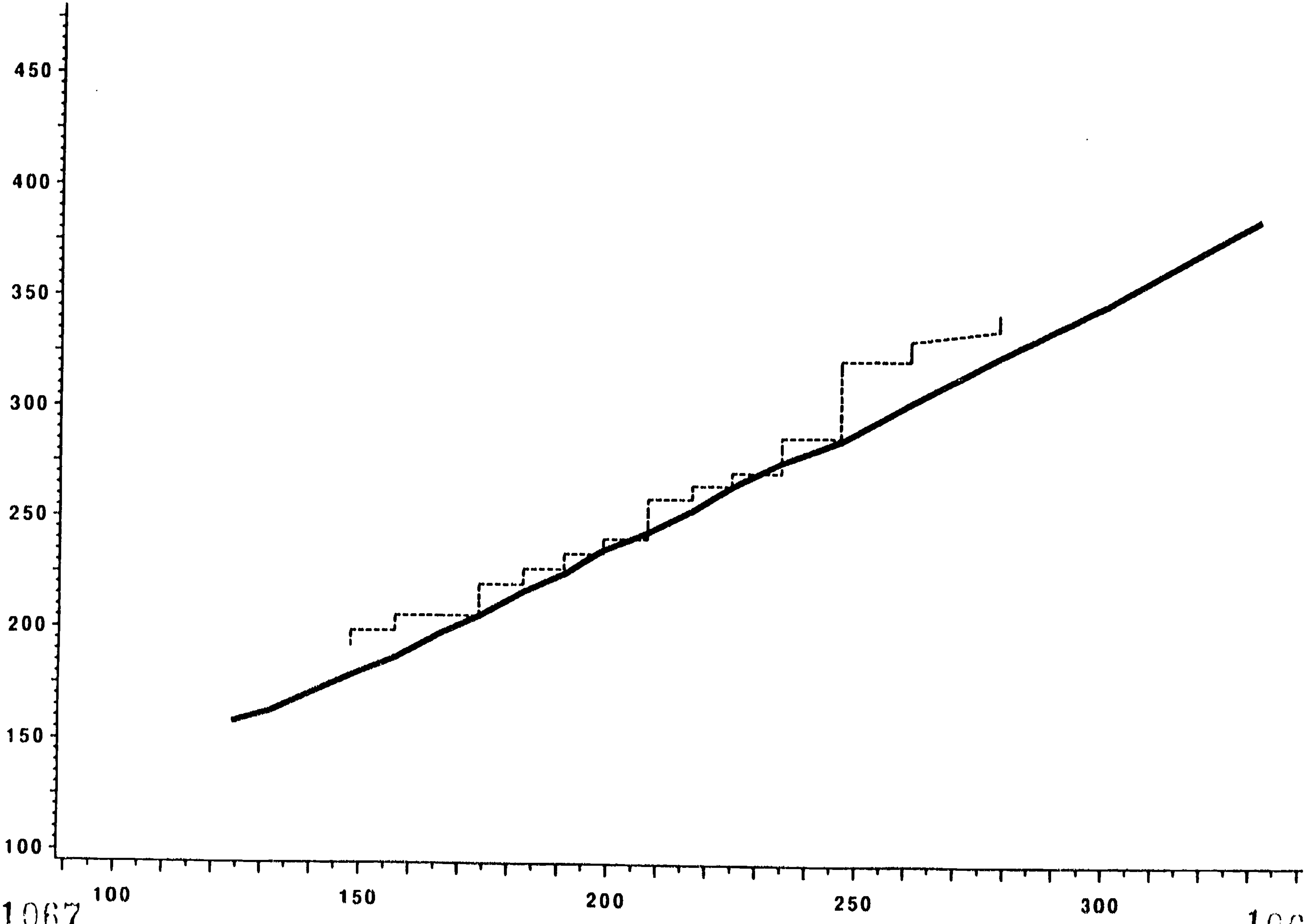
- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 88

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



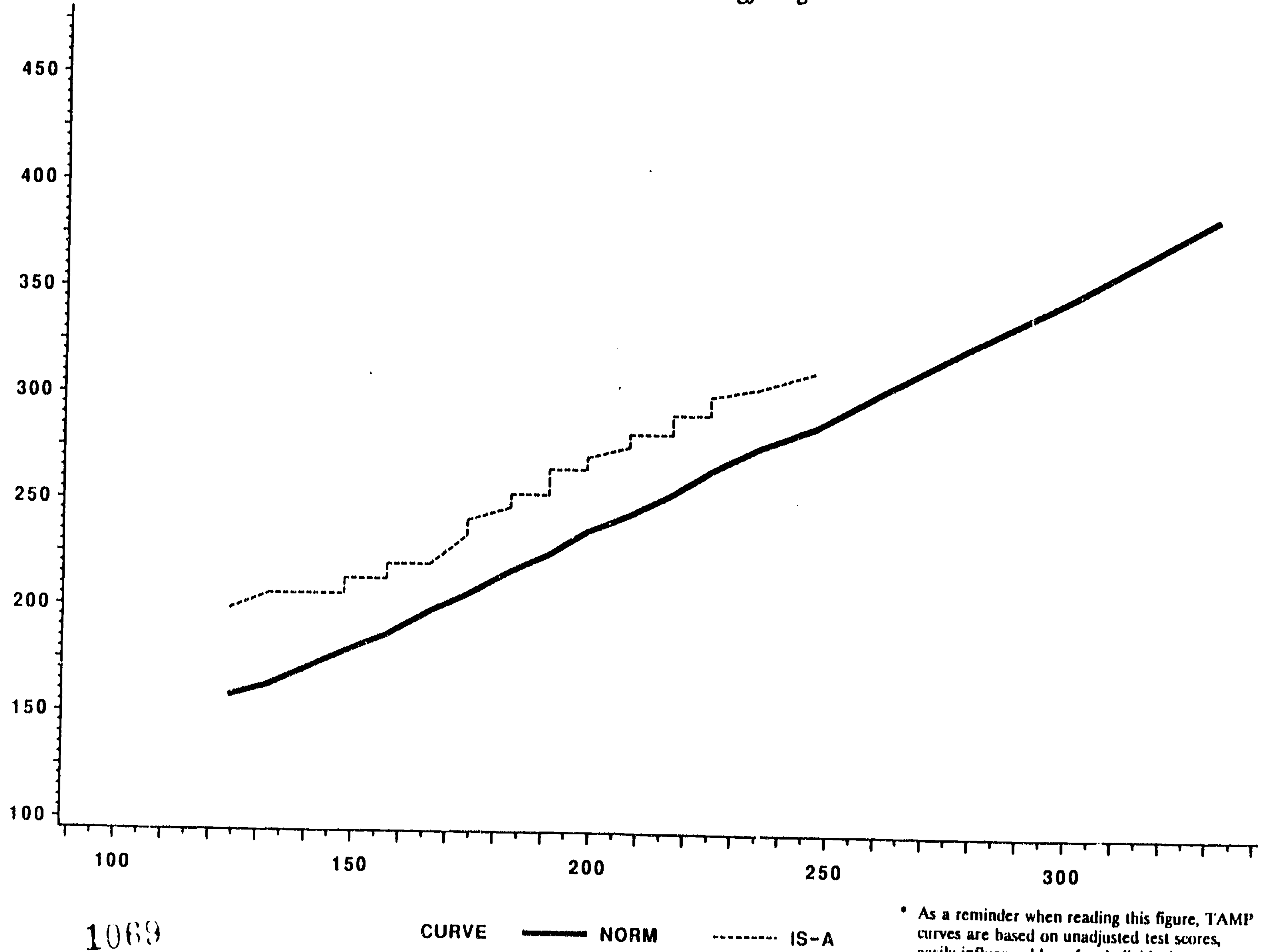
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

Figure 89

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

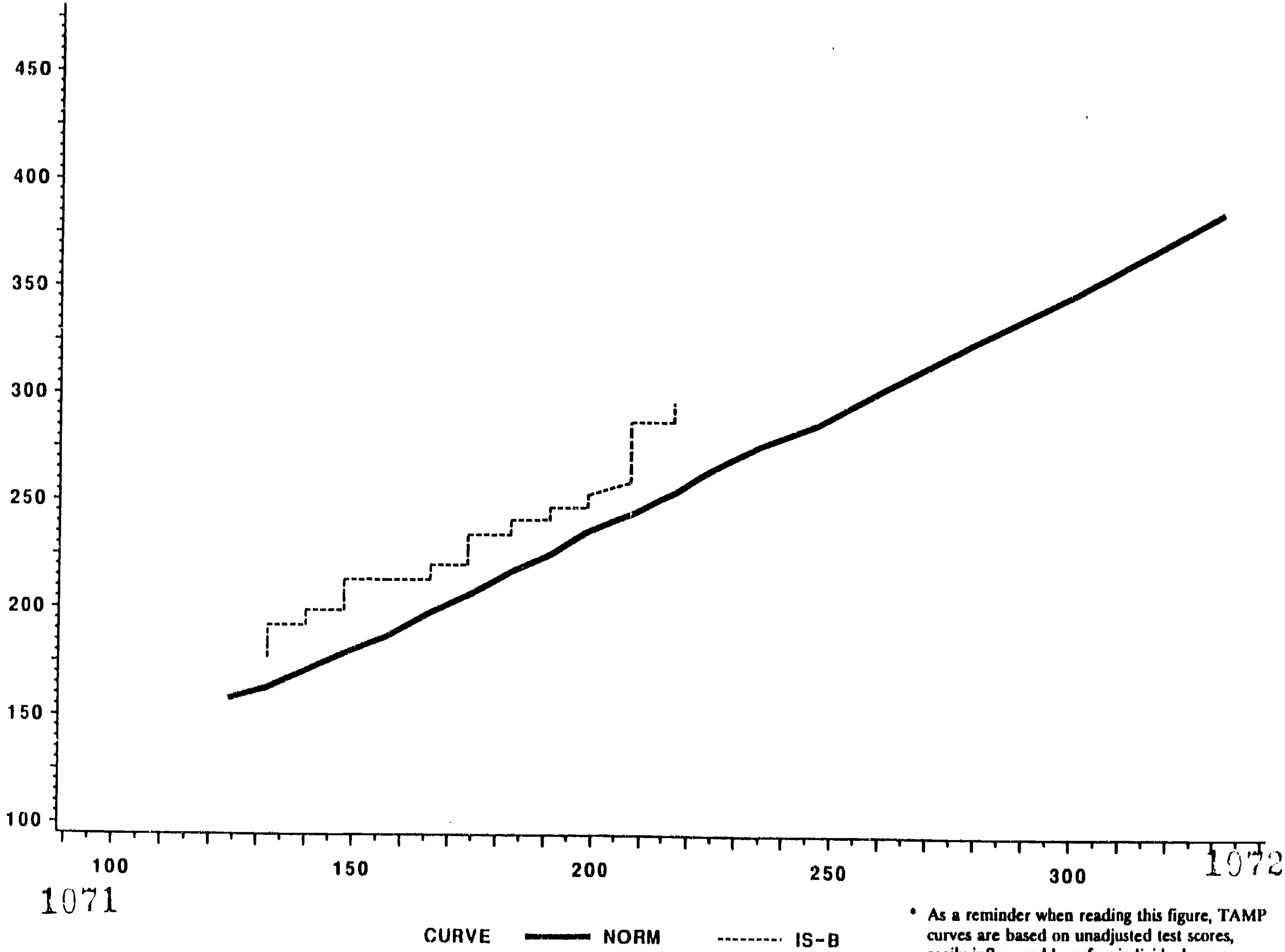
453

ERIC Full Text Provided by ERIC

Figure 90

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



454

1071

1072

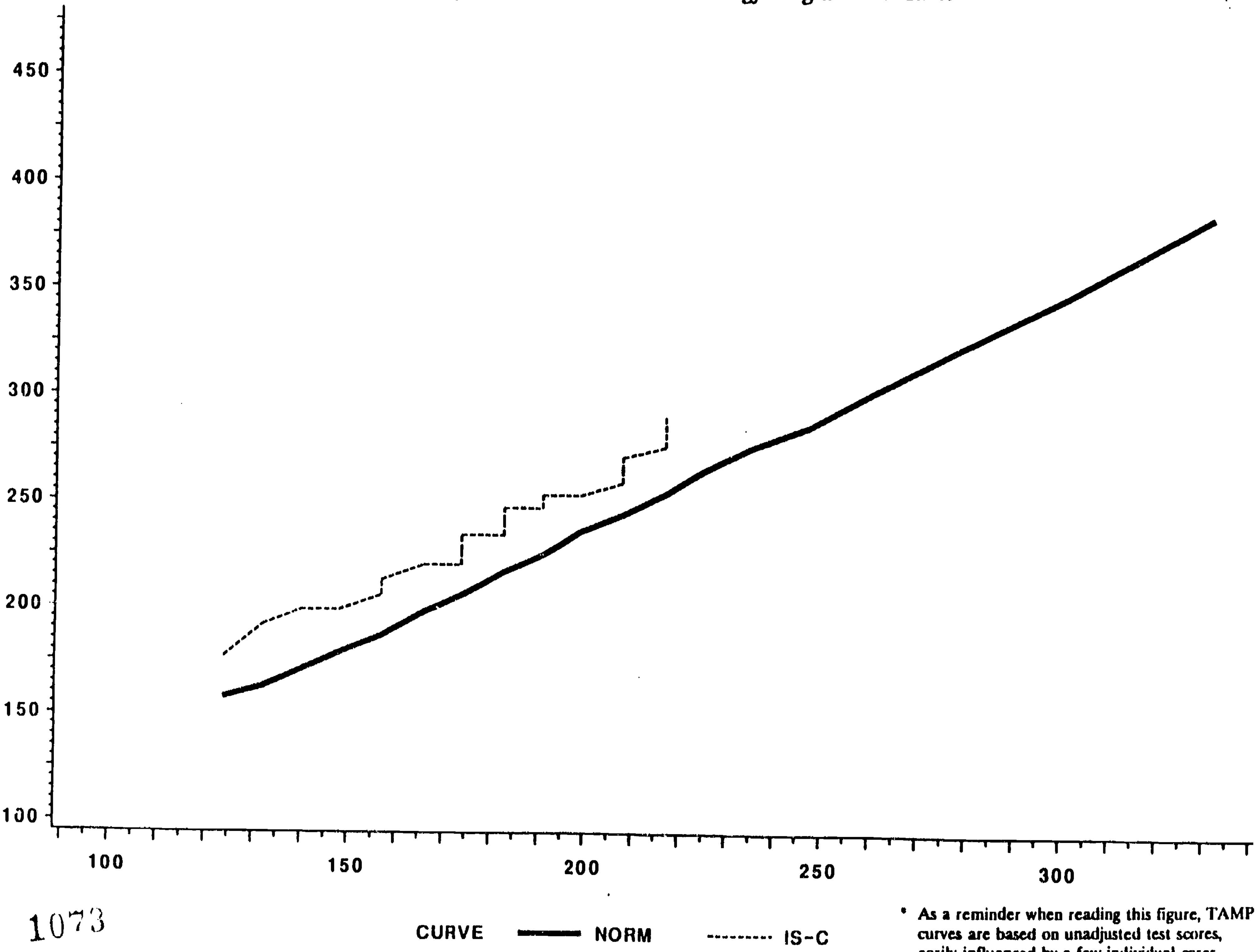
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 91

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



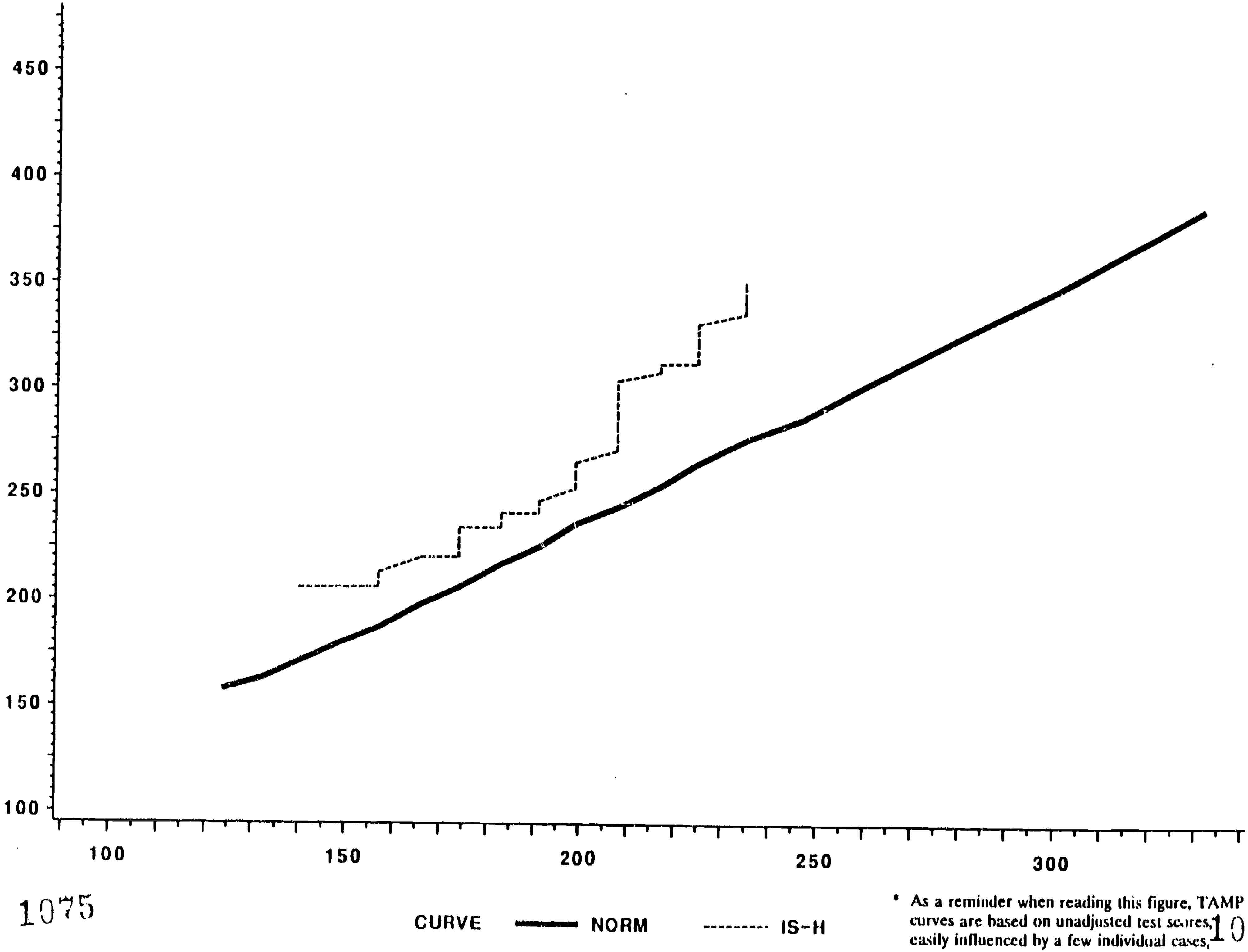
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 92

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H



456

1075

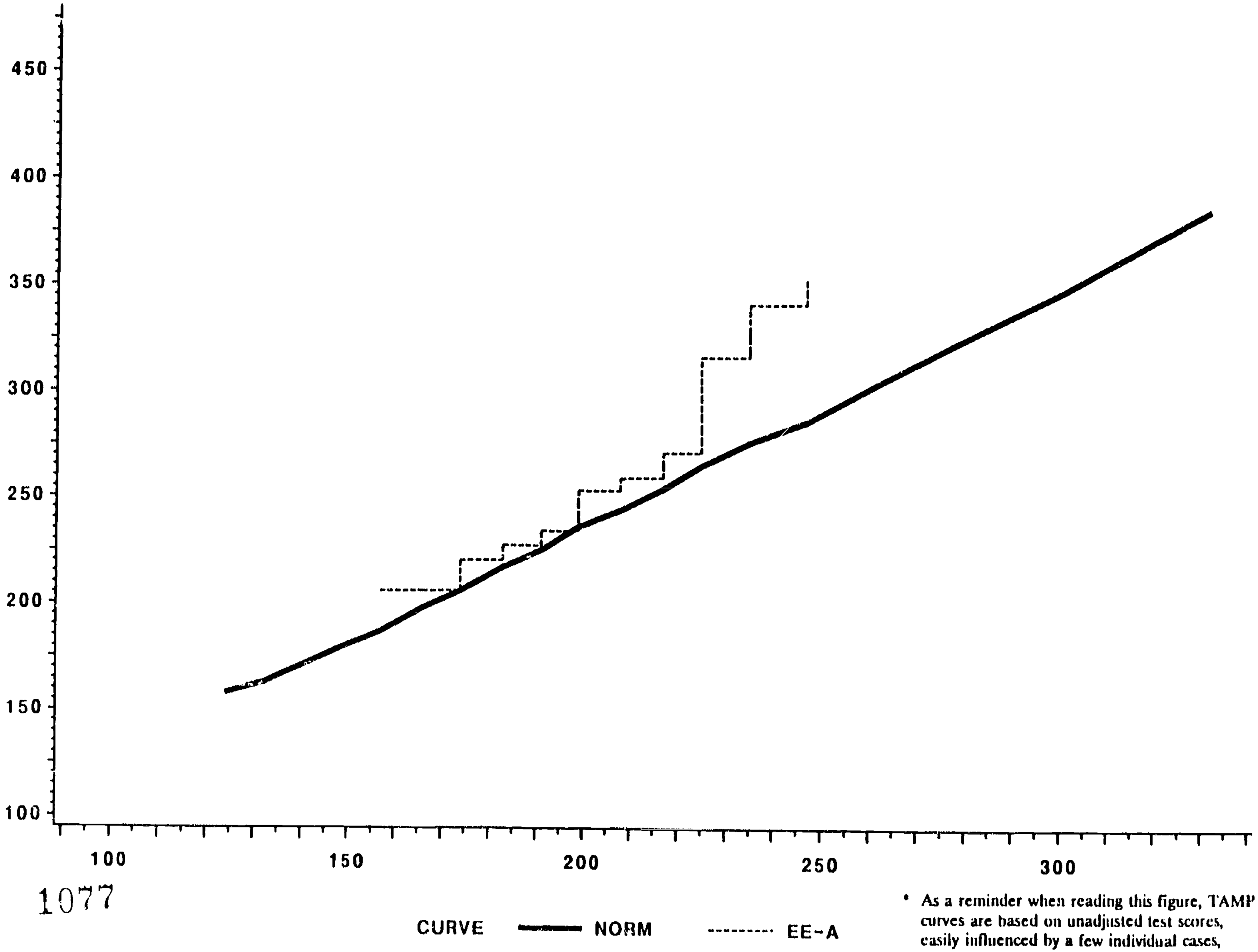
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

1076

Figure 93

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A

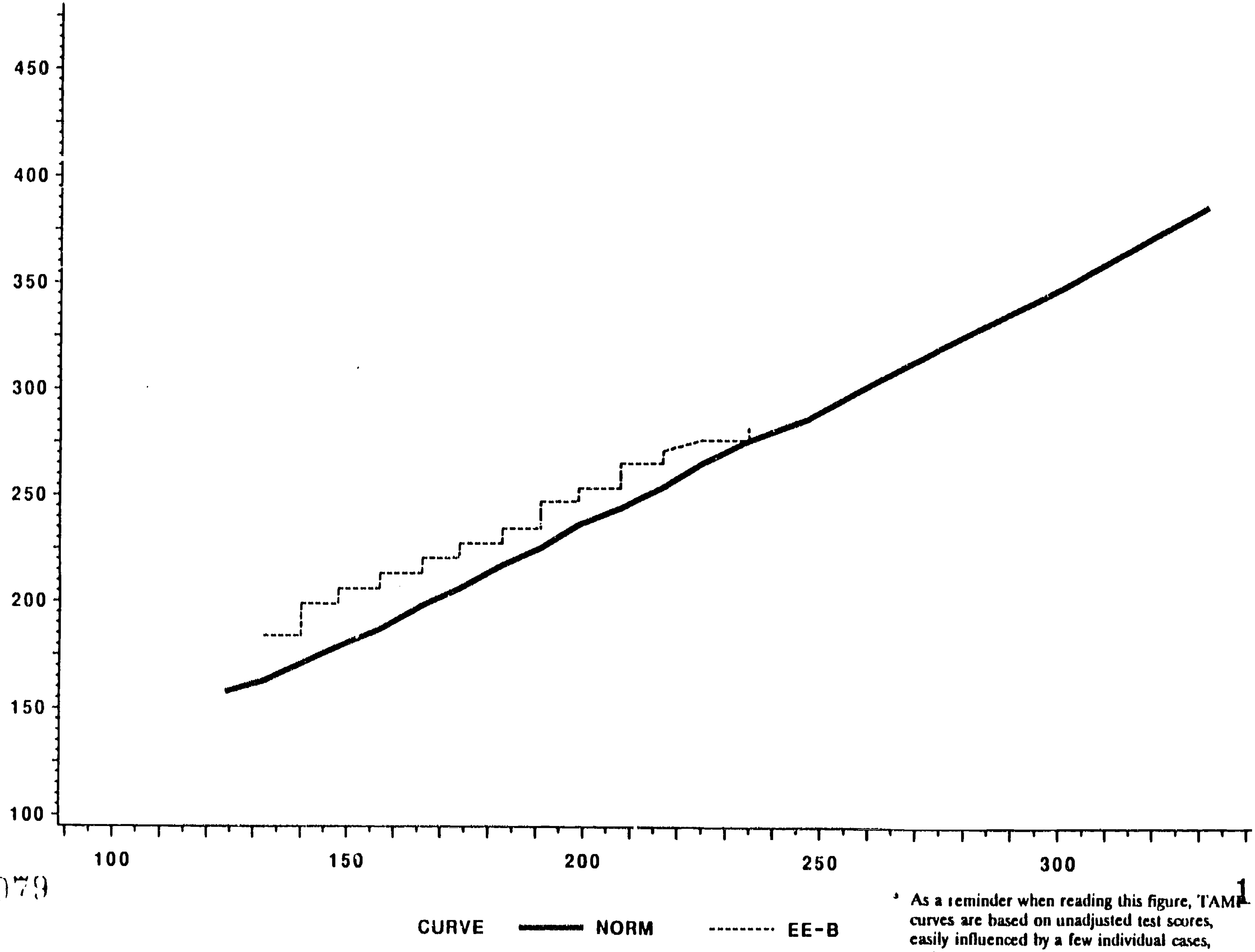


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 94

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



458

1079

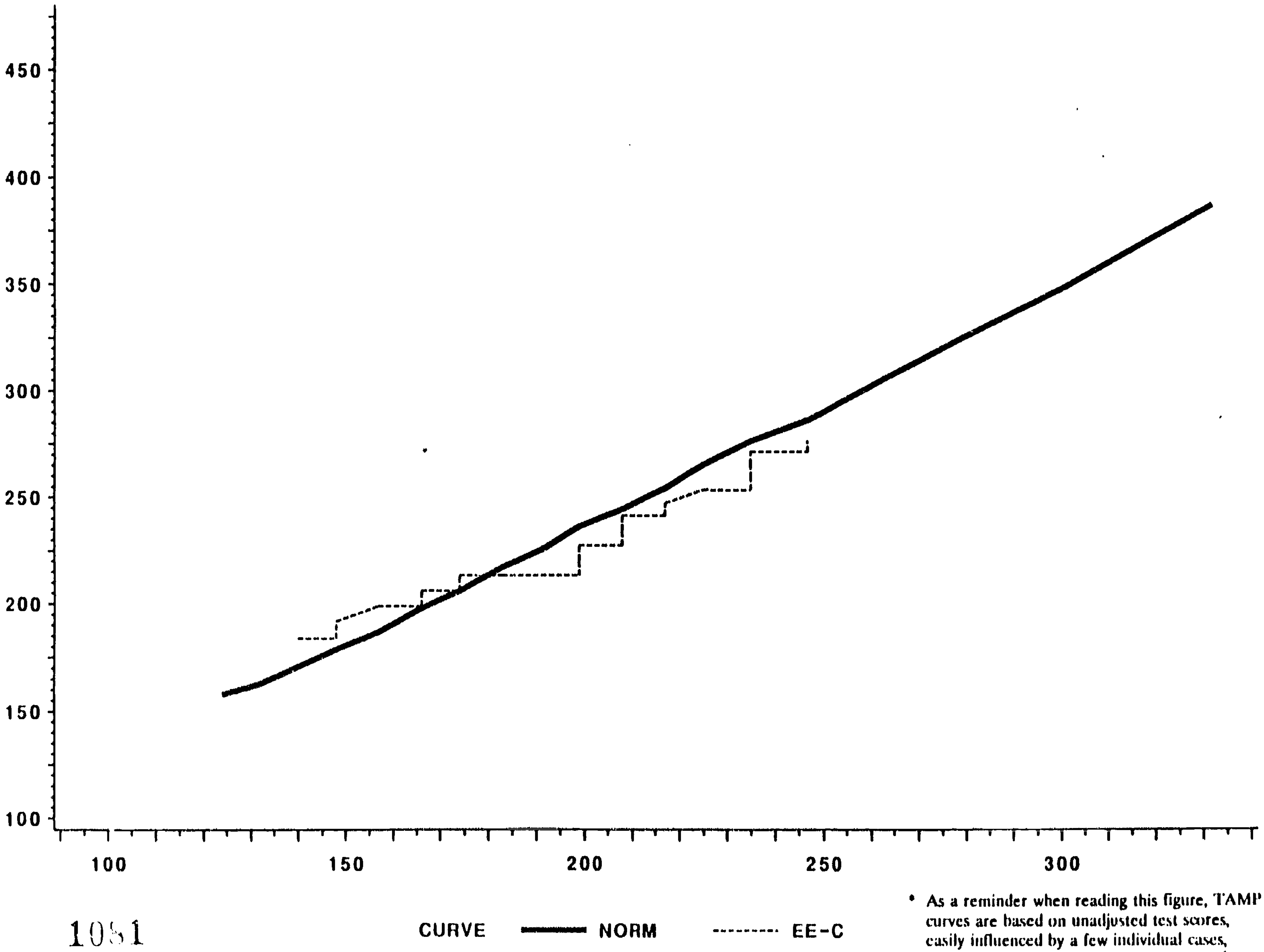
1080

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 95

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



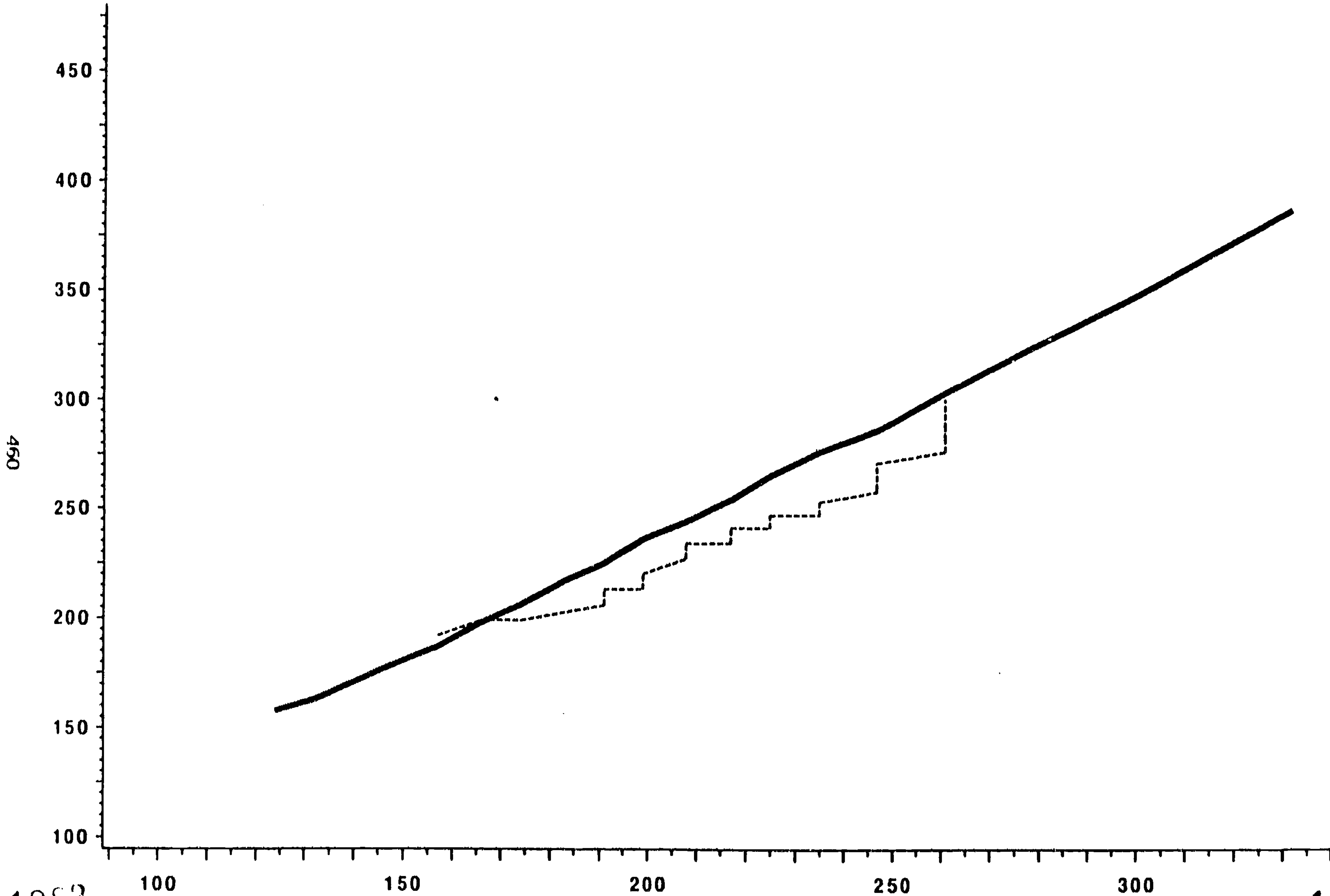
459

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 96

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores,

D.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Language**	**to**	**Language**

Immersion strategy, early-exit, and late-exit students all appeared to grow at a faster rate than this norming population in the area of English language skills (see Figures 97, 98, and 99). This bodes well for all three programs in their efforts to help their students approximate the achievement of this norming population. This is surprising as, given their respective instructional models, late-exit students were expected to show only minimal growth given their limited instruction in English. Noteworthy is that the upward jog in growth rate that immersion strategy students seemed to experience in mathematics was not replicated with English language skills. That is, the lower Spanish mathematics skills in spring kindergarten for immersion strategy students did not show up when mathematics was tested in English. The difference may reflect that mathematics skills at these lower grade levels do not rely as much upon verbal input and/or are more amenable to the use of special methodology (e.g., use of realia) to support instruction in English. As a result, greater growth would be expected in mathematics than in English language skills.

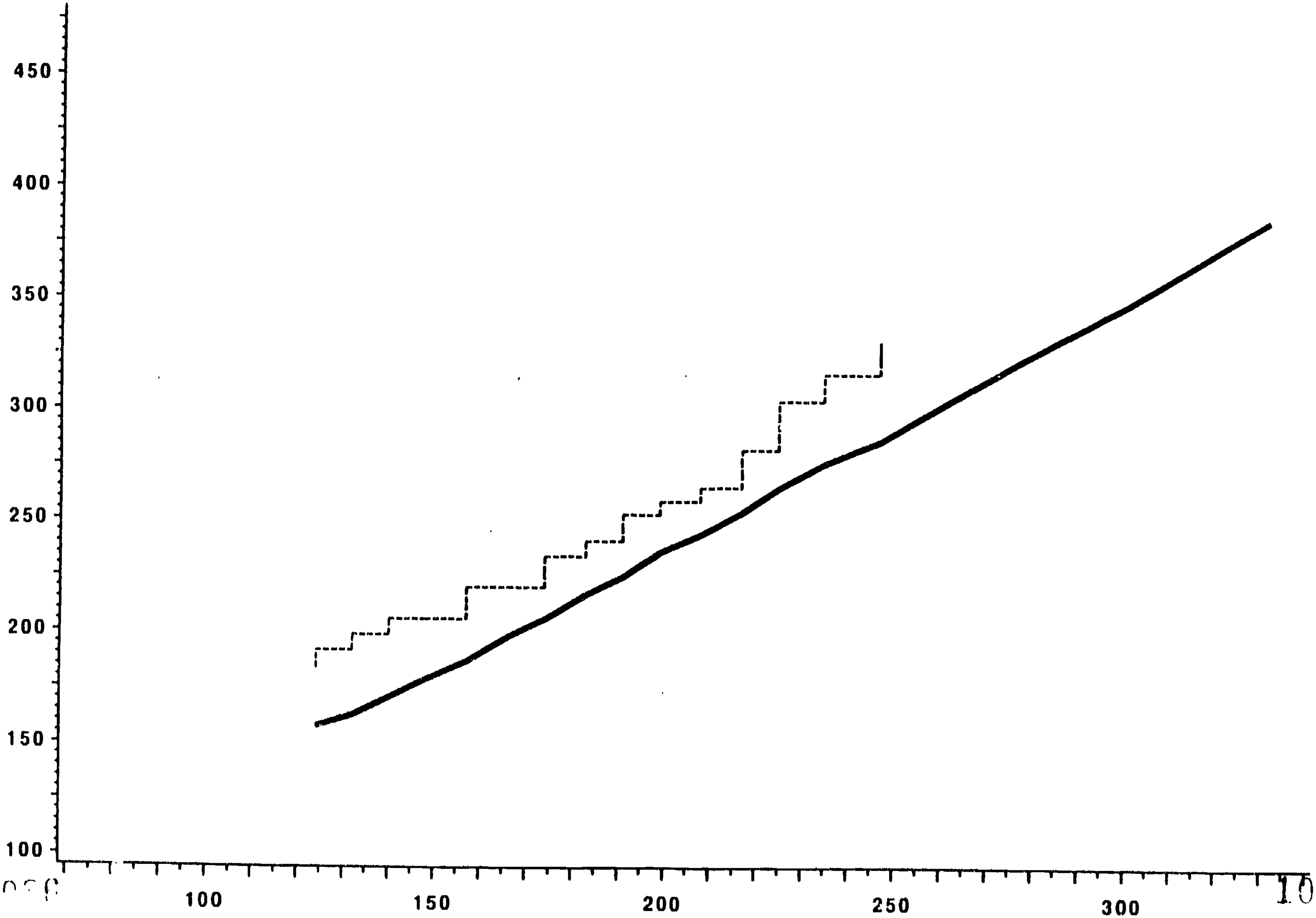
No differences in growth were noted between the individual immersion strategy sites (see Figures 100 to 103). While the four early-exit sites also appear to have grown at the same rate, clearly site EE-A students had higher initial English language skills than in either site EE-B or EE-C, and students in site EE-H grew the slowest (see Figures 104 to 107). In fact, the distribution of spring first grade English language scores in site EE-H is scarcely related to the spring kindergarten scores; the TAMP curve is nearly horizontal, indicating a marked narrowing of the distribution of the test scores.

ERIC

Figure 97

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

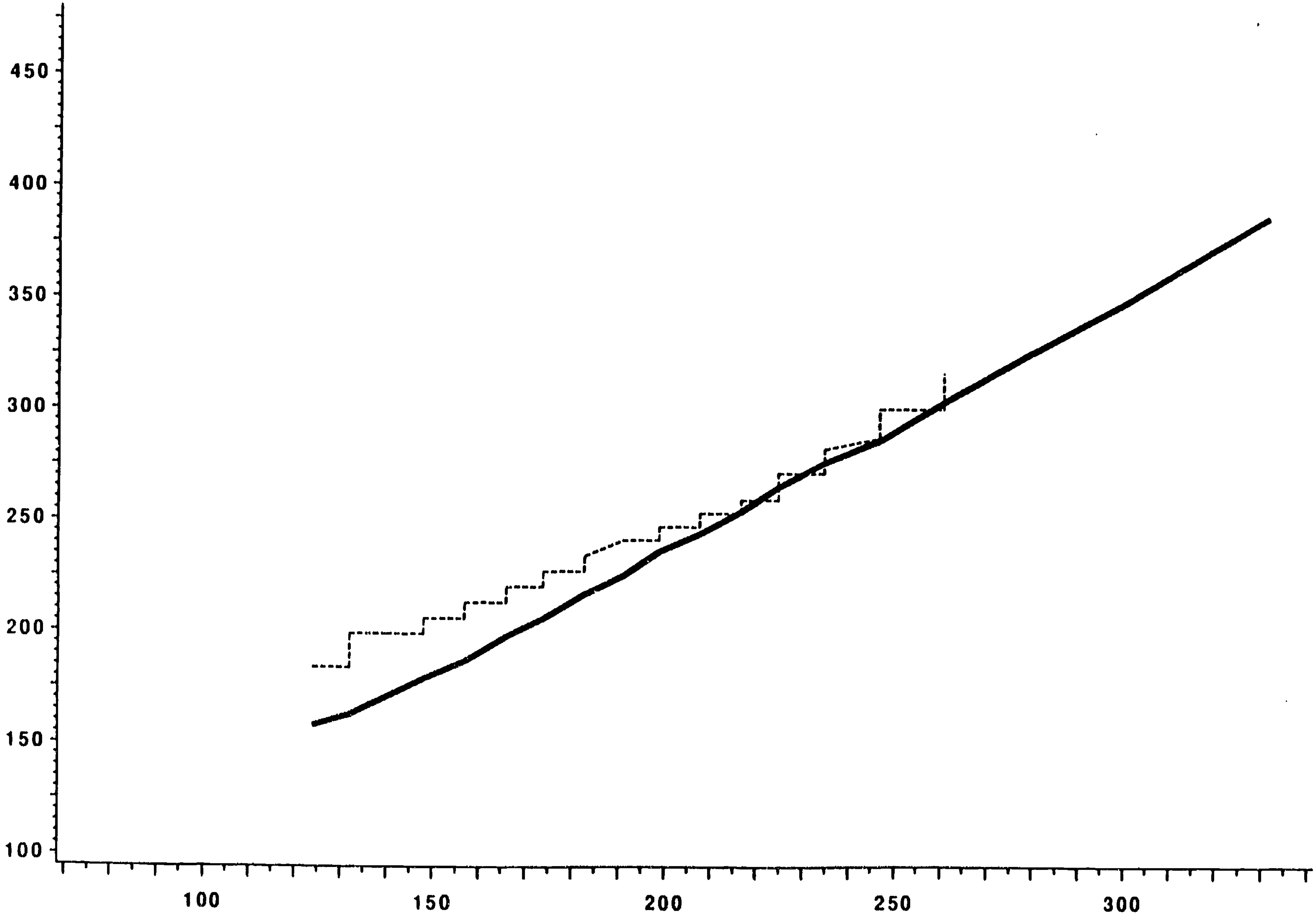


* As a reminder when reading this figure, TAMP

Figure 98

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

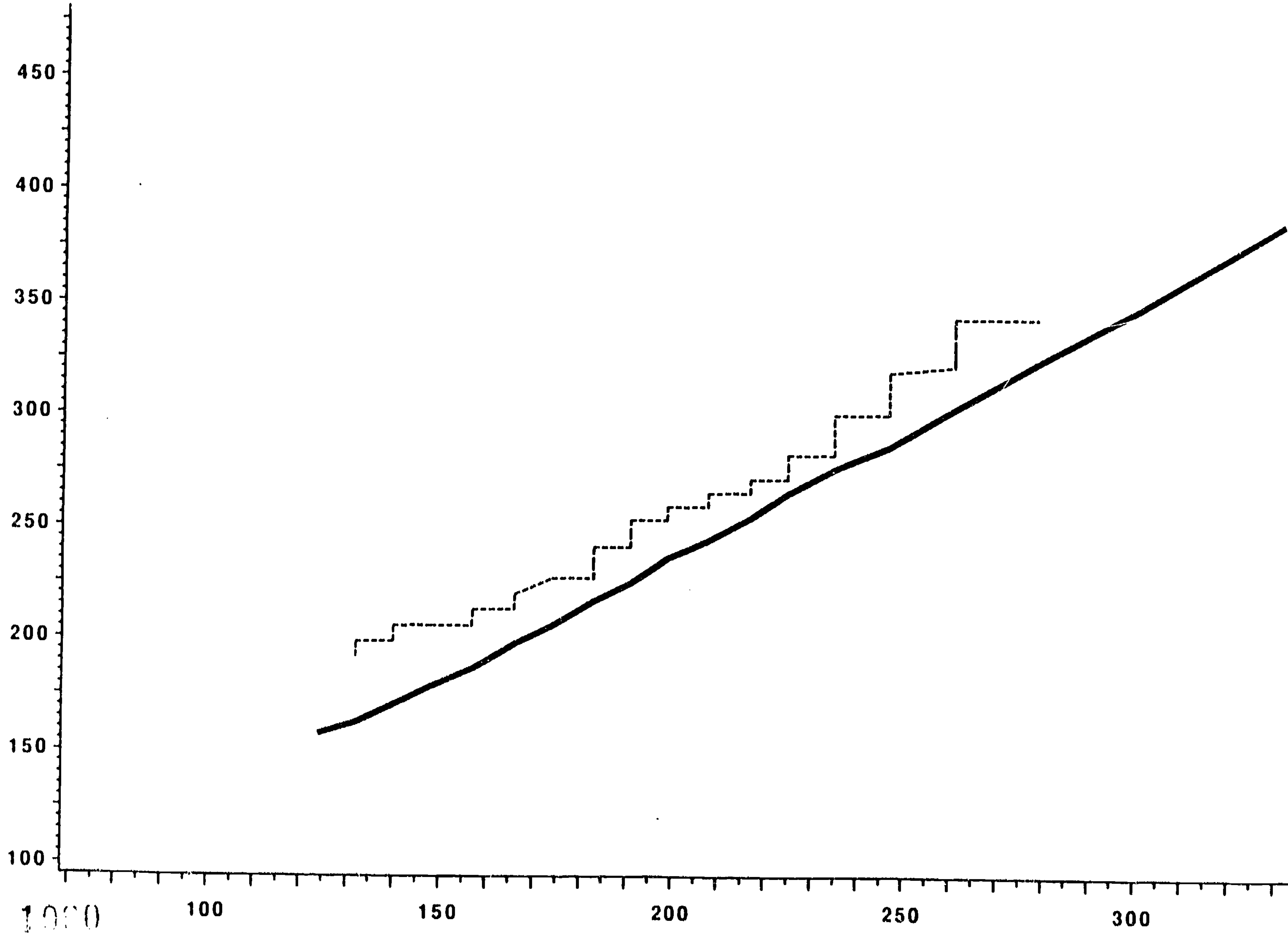


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 99

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



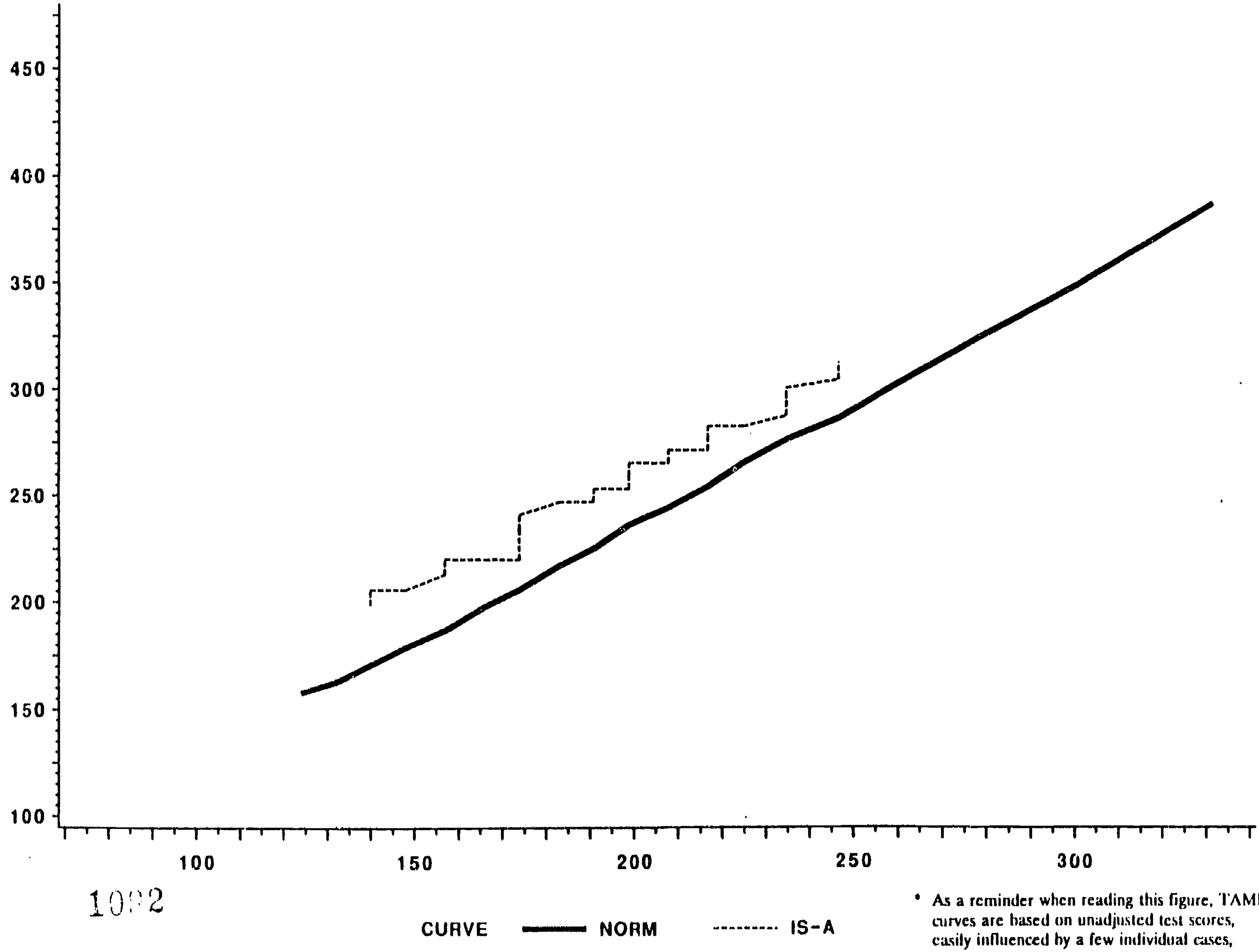
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores.

1090 1091

Figure 100

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A

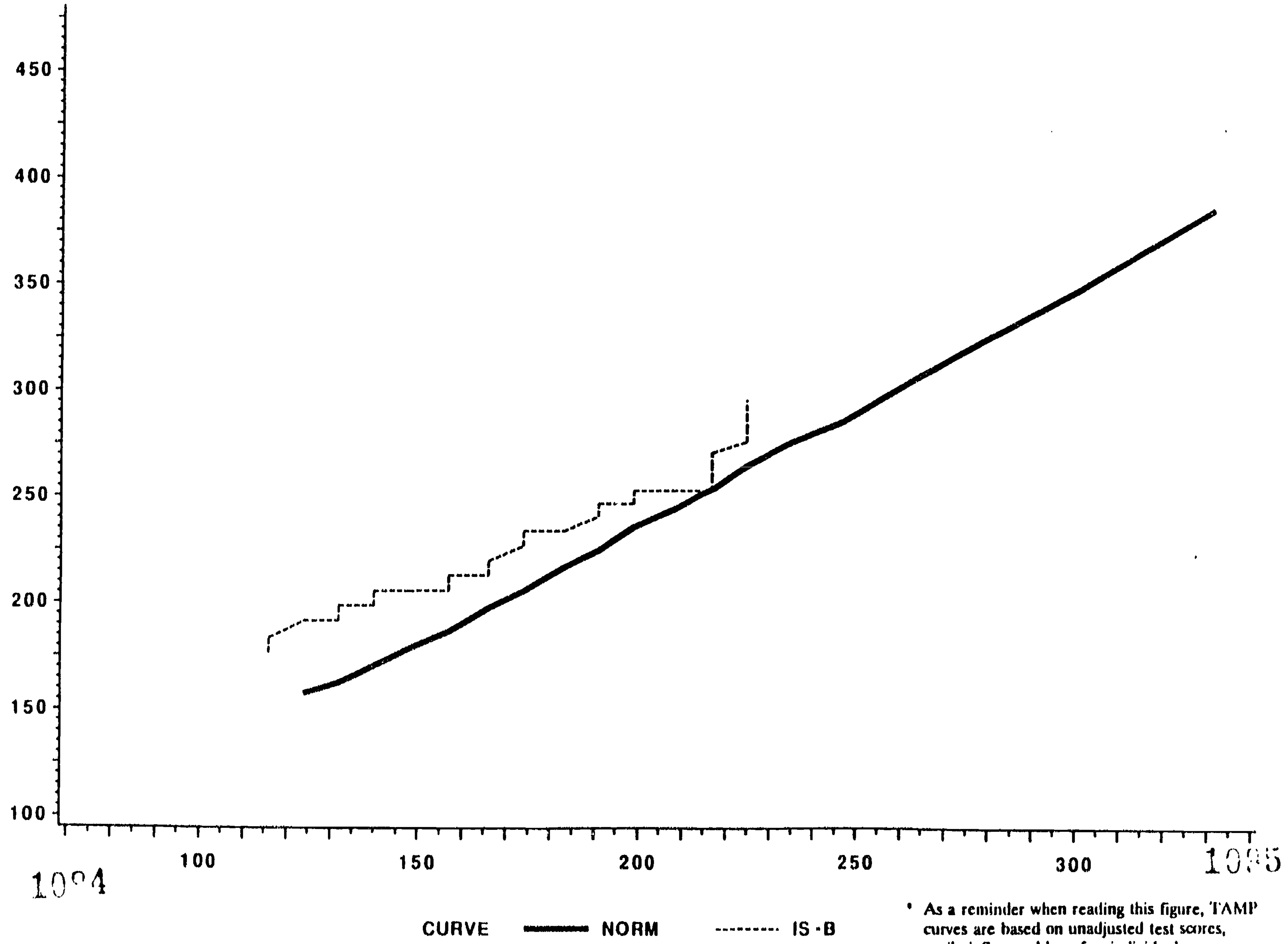


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 101

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



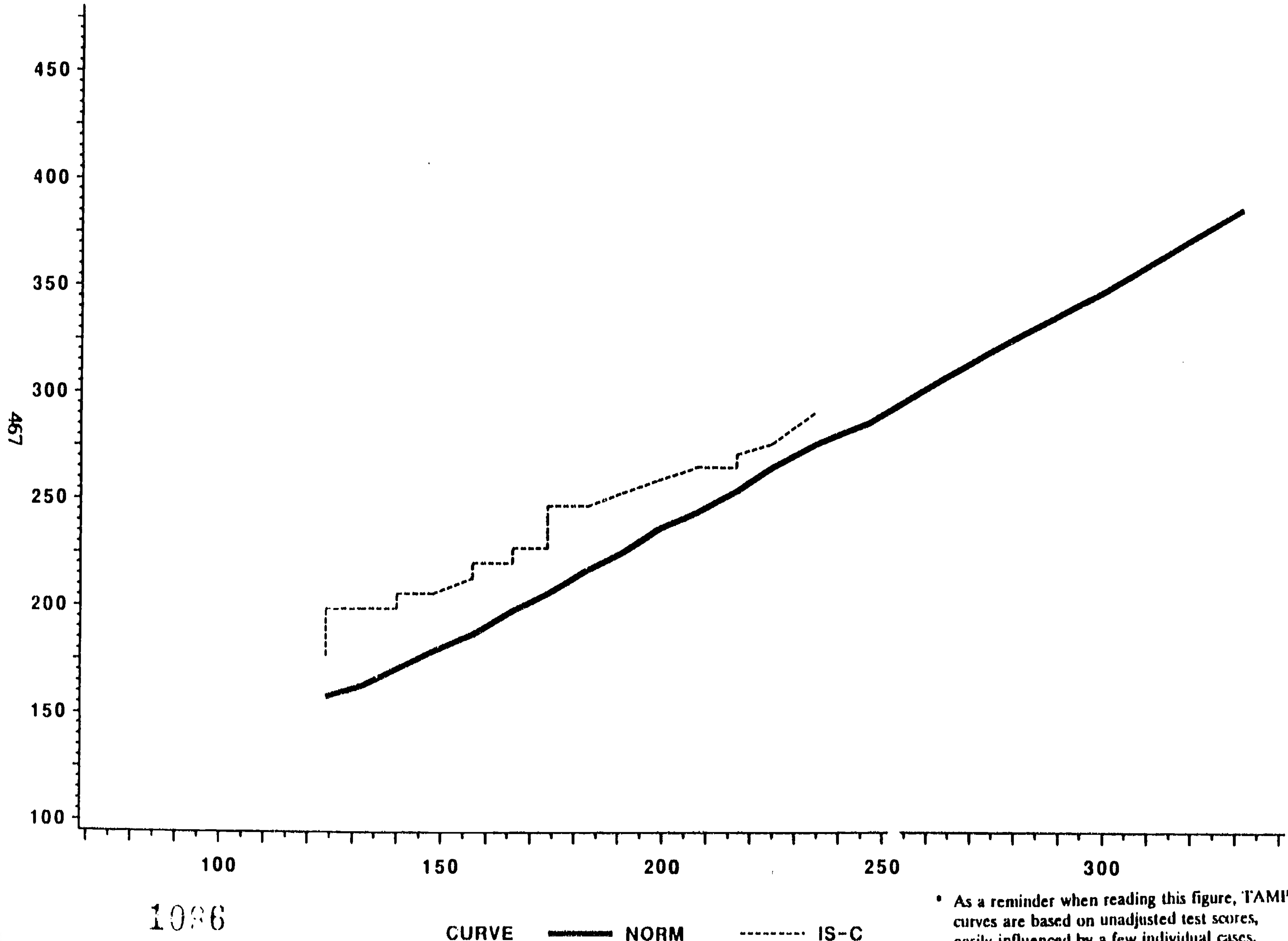
466

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 102

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



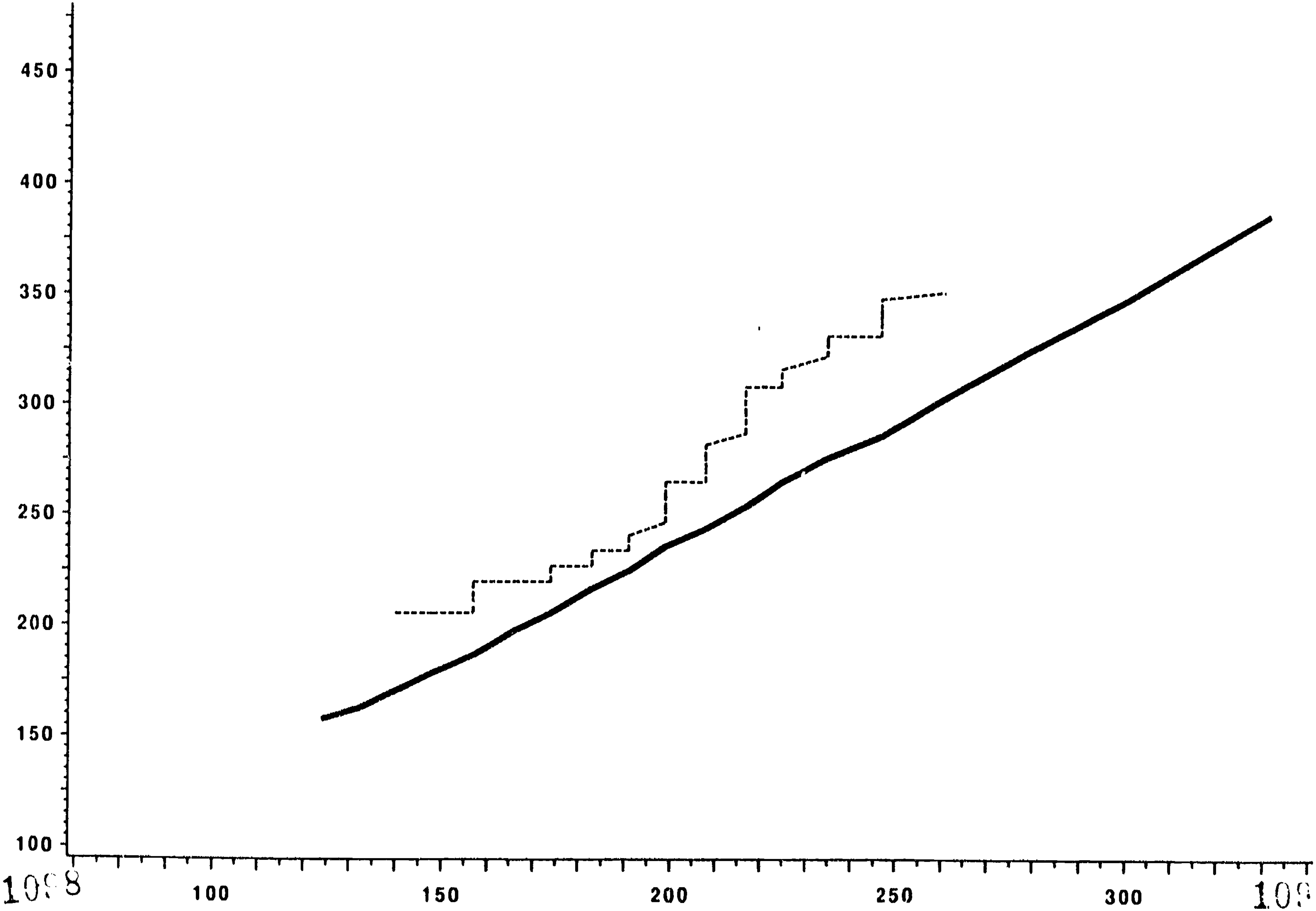
- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 103

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H



468

1098 1099

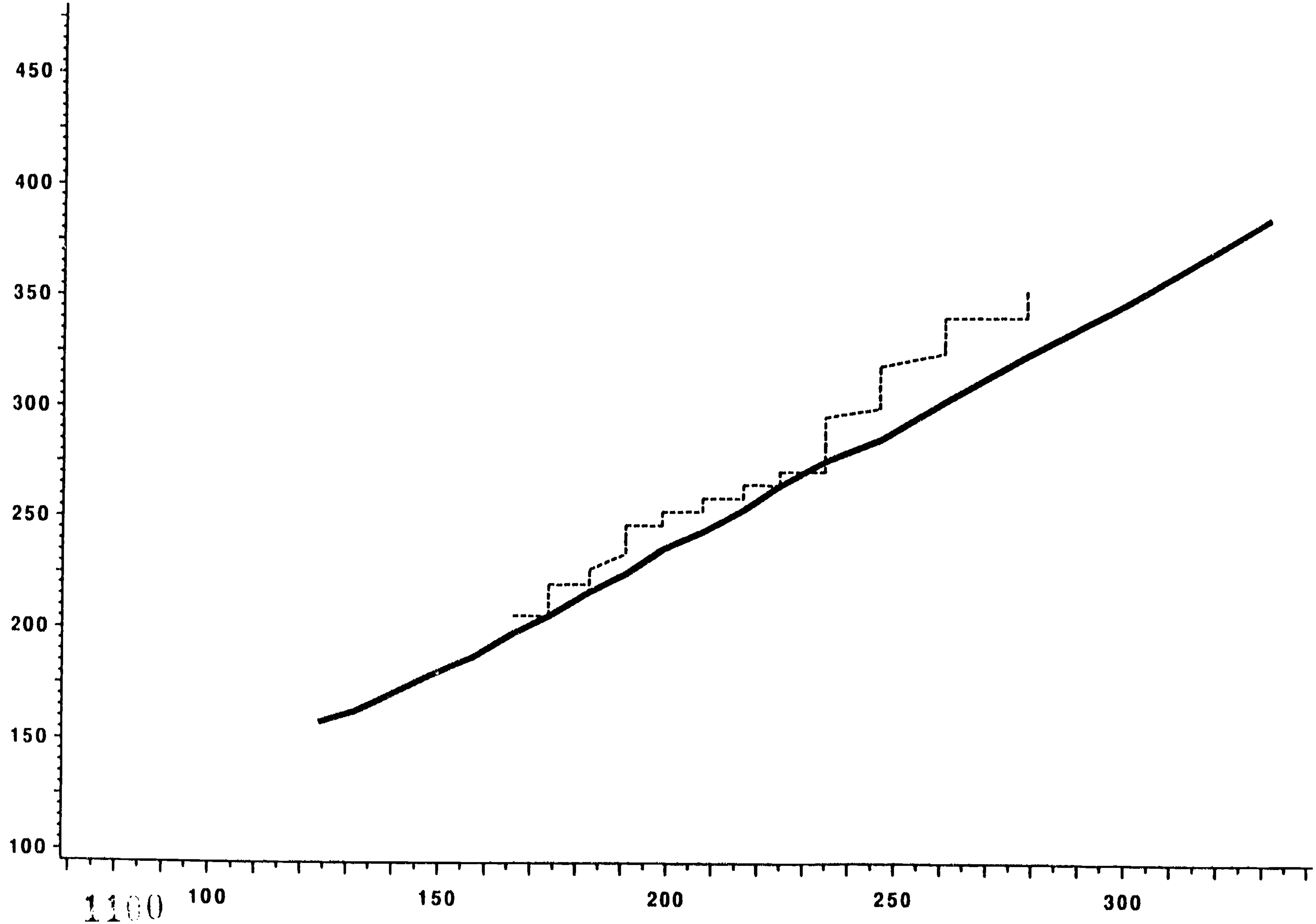
* As a reminder when reading this figure, TAMP

ERIC

Figure 104

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A

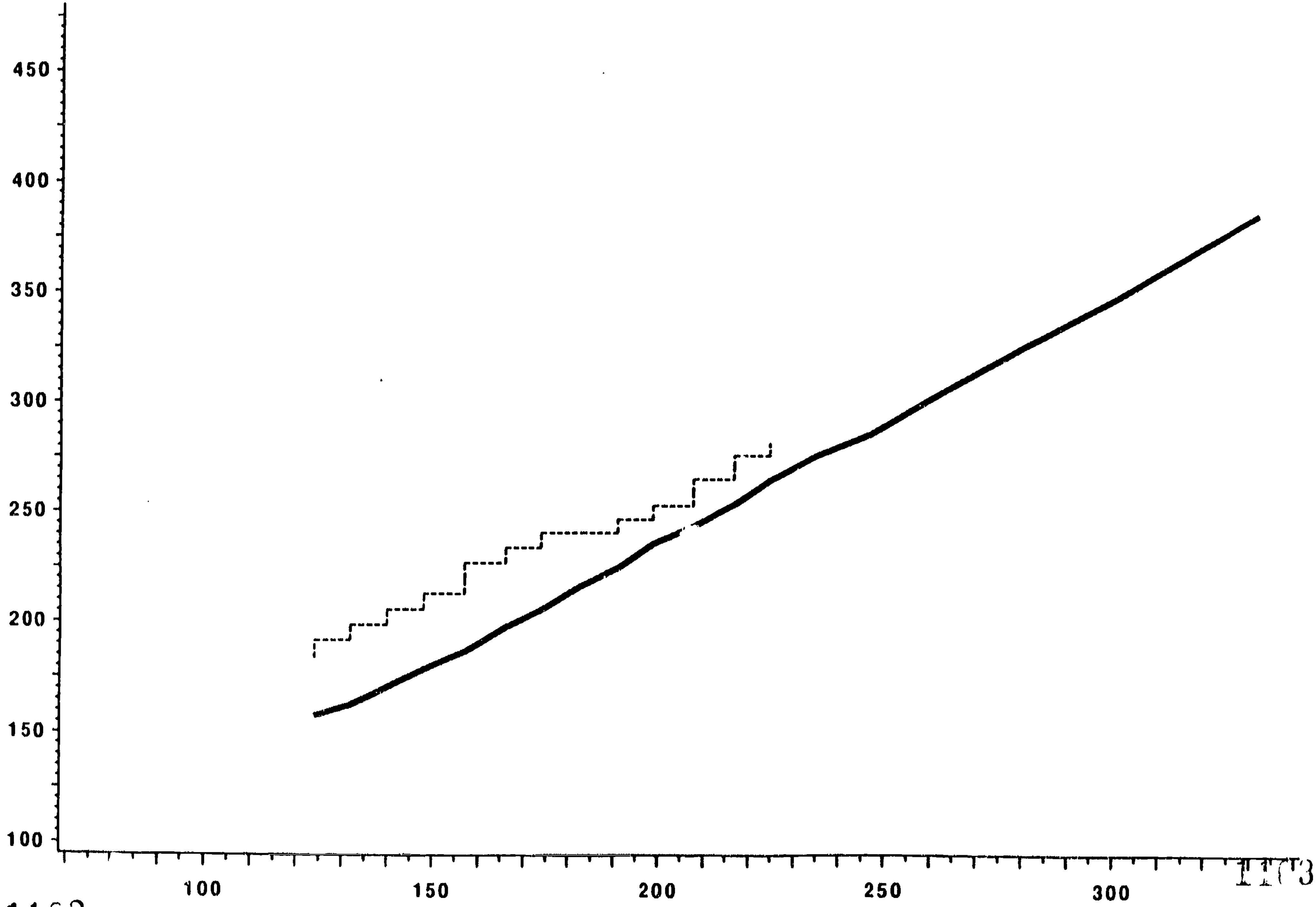


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 105

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B

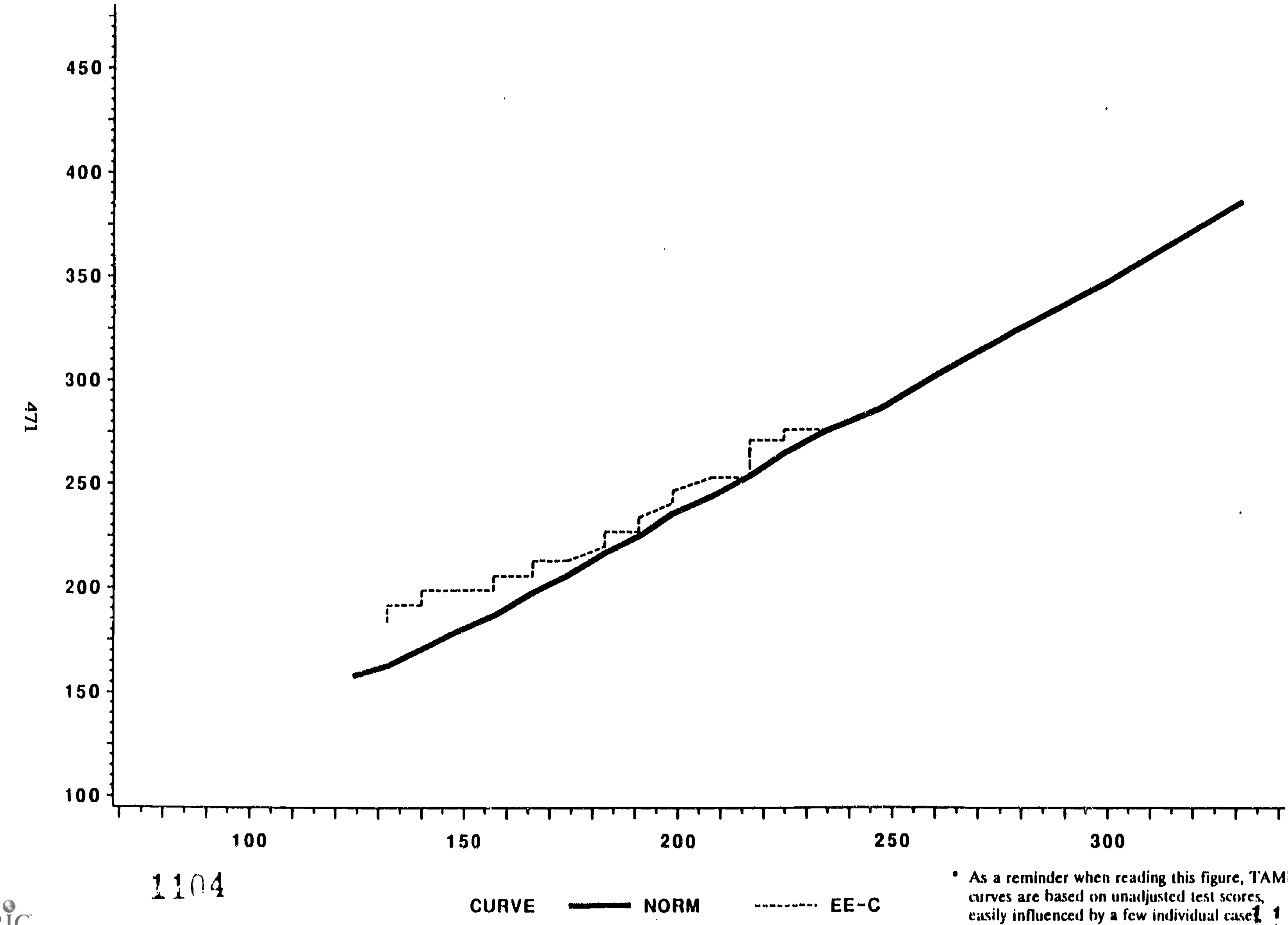


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 106

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



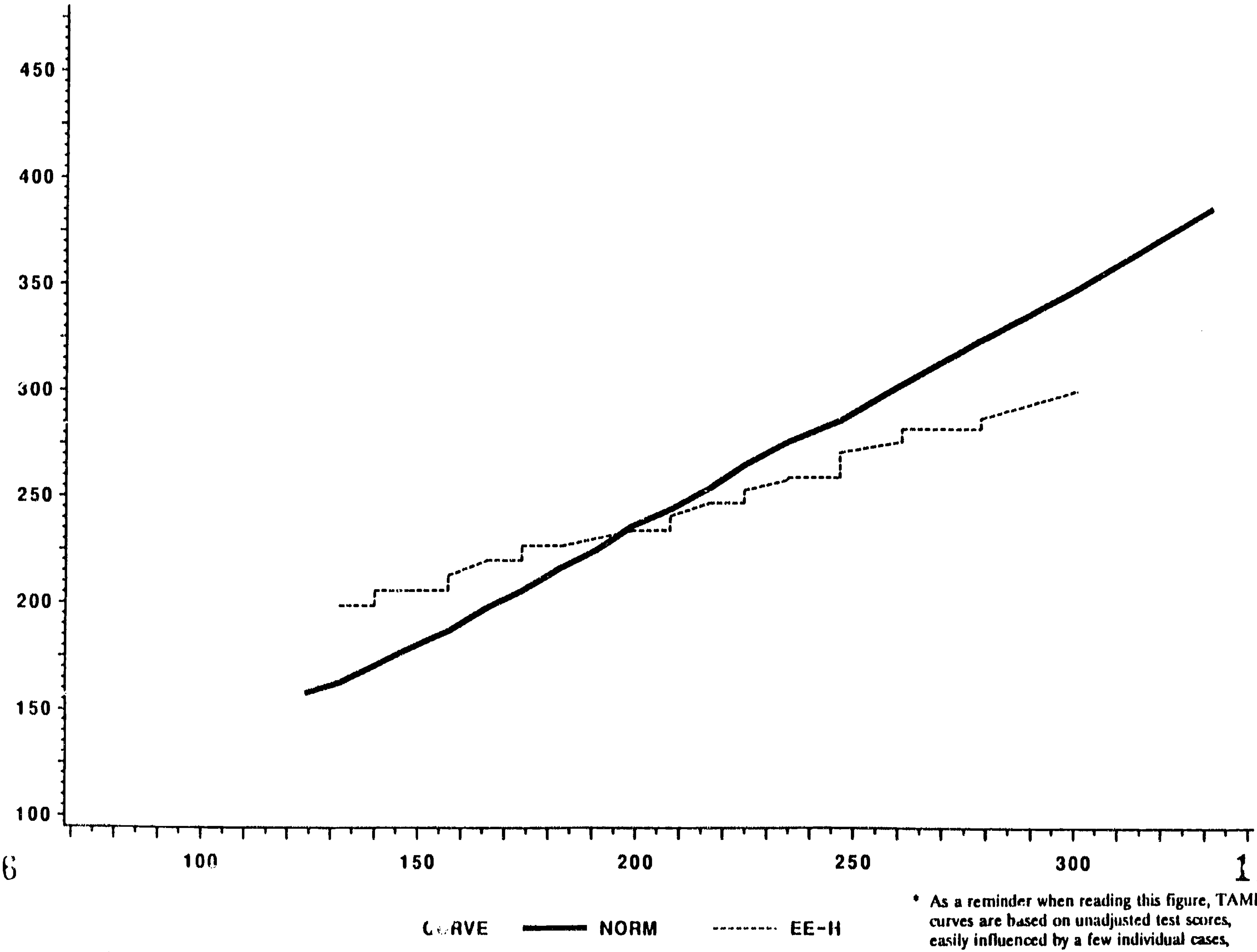
1104

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 107

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and a[illegible]bject [illegible]mplin[illegible]uatio[illegible]

ERIC

E.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Language**	**to**	**Reading**

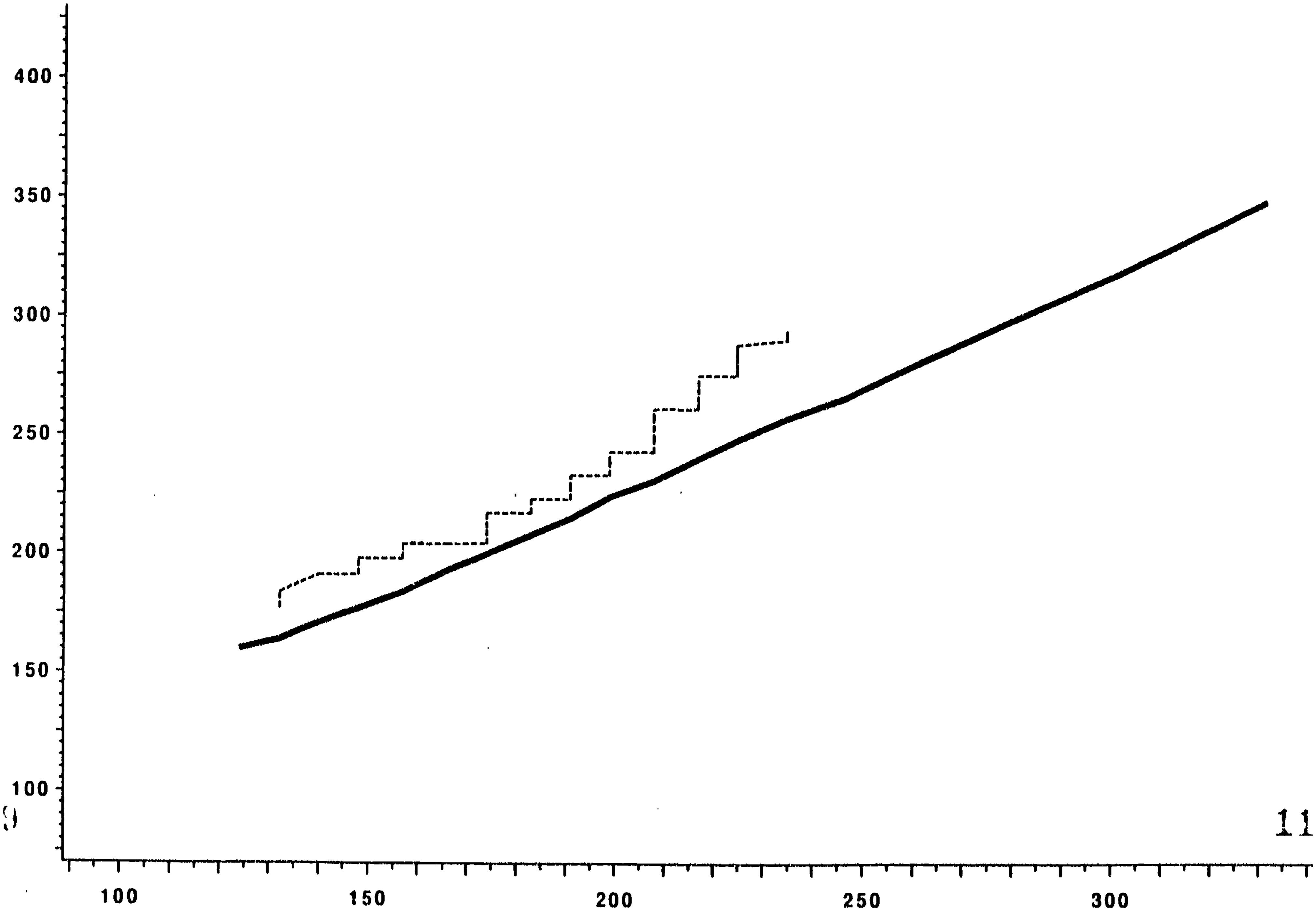
As above, these analyses are concerned with determining the extent to which immersion strategy, early-exit, and late-exit students ended kindergarten with comparable Spanish language skills and comparable growth rates in English reading skills to one another and to this norming population. The TAMP curves indicate that students in all three programs began with roughly the same entry-level Spanish language skills and exhibited the same growth rate as each other (see Figures 108, 109, and 110). It appears that students in each of the three programs exhibited a faster increase in English reading skills relative to this norming population. There is a slight indication that immersion strategy students exhibited a consistently higher growth rate in English reading relative to this norming population, early-exit, and late-exit students. While there was no variation in academic growth between immersion strategy sites (see Figures 111 to 114), two of the four early-exit sites seem to be slightly different from the other two sites. It seems that students in site EE-A with the highest initial skills learned at a much faster rate than this norming population and comparable students in the other three early-exit sites, and site EE-H students grew more slowly than this norming population and the other early-exit sites (see Figures 115 to 118).

In sum, consistently across the three programs students appear to have improved at a rate that was equal to, and often better than, this norming population. As predicted, the initial scores and the growth rates in the TAMP curves for immersion strategy and early-exit students appear to be about the same. Contrary to expectations, even though late-exit students had limited instruction in English, they learned to read faster than this norming population relative to where they stood at spring of kindergarten. That is, the late-exit students were catching up after losing ground from fall kindergarten to spring kindergarten. Finally, it seems that late-exit students grew as fast as immersion strategy students.

Figure 108

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



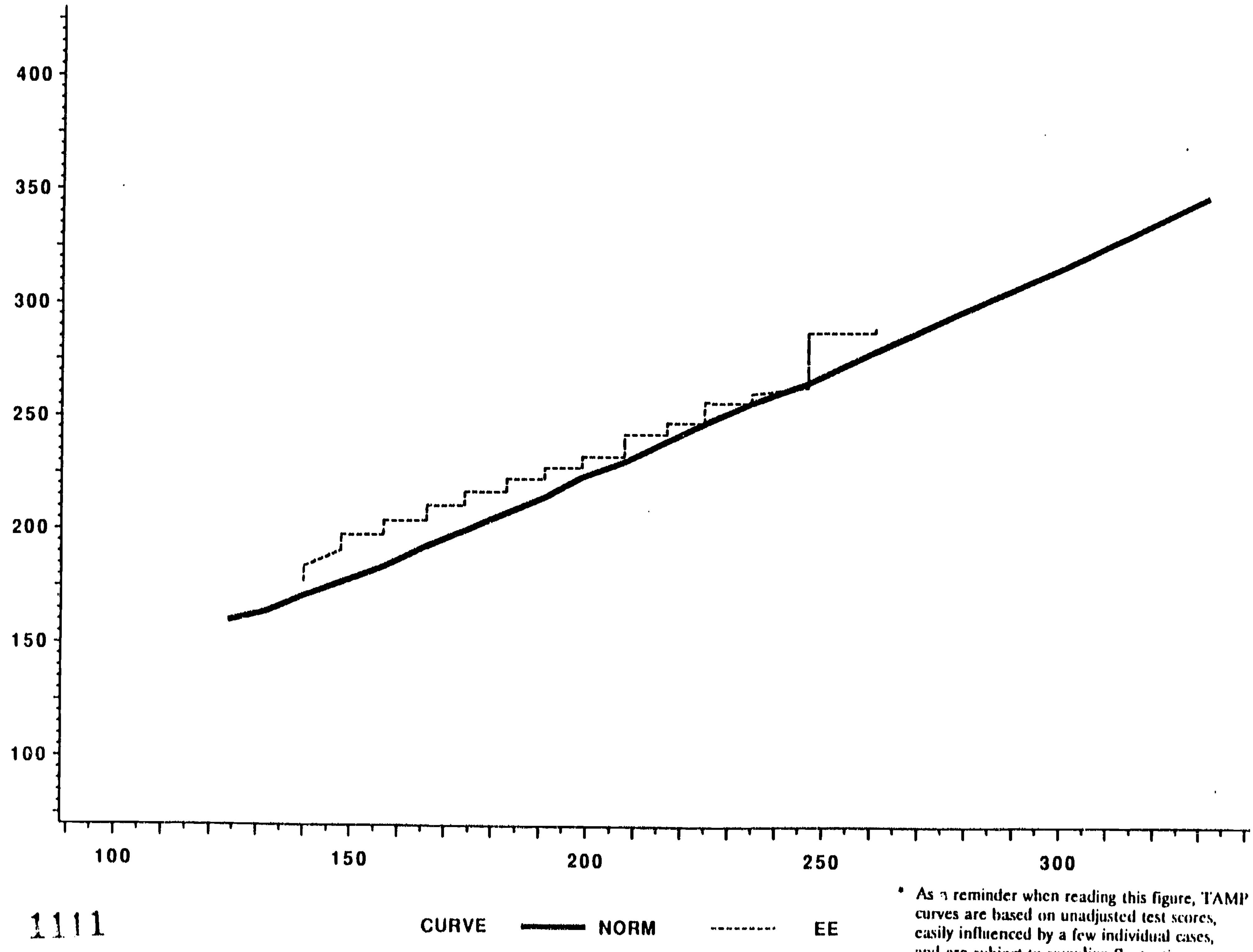
ERIC

* As a reminder when reading this figure, TAMP

Figure 109

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

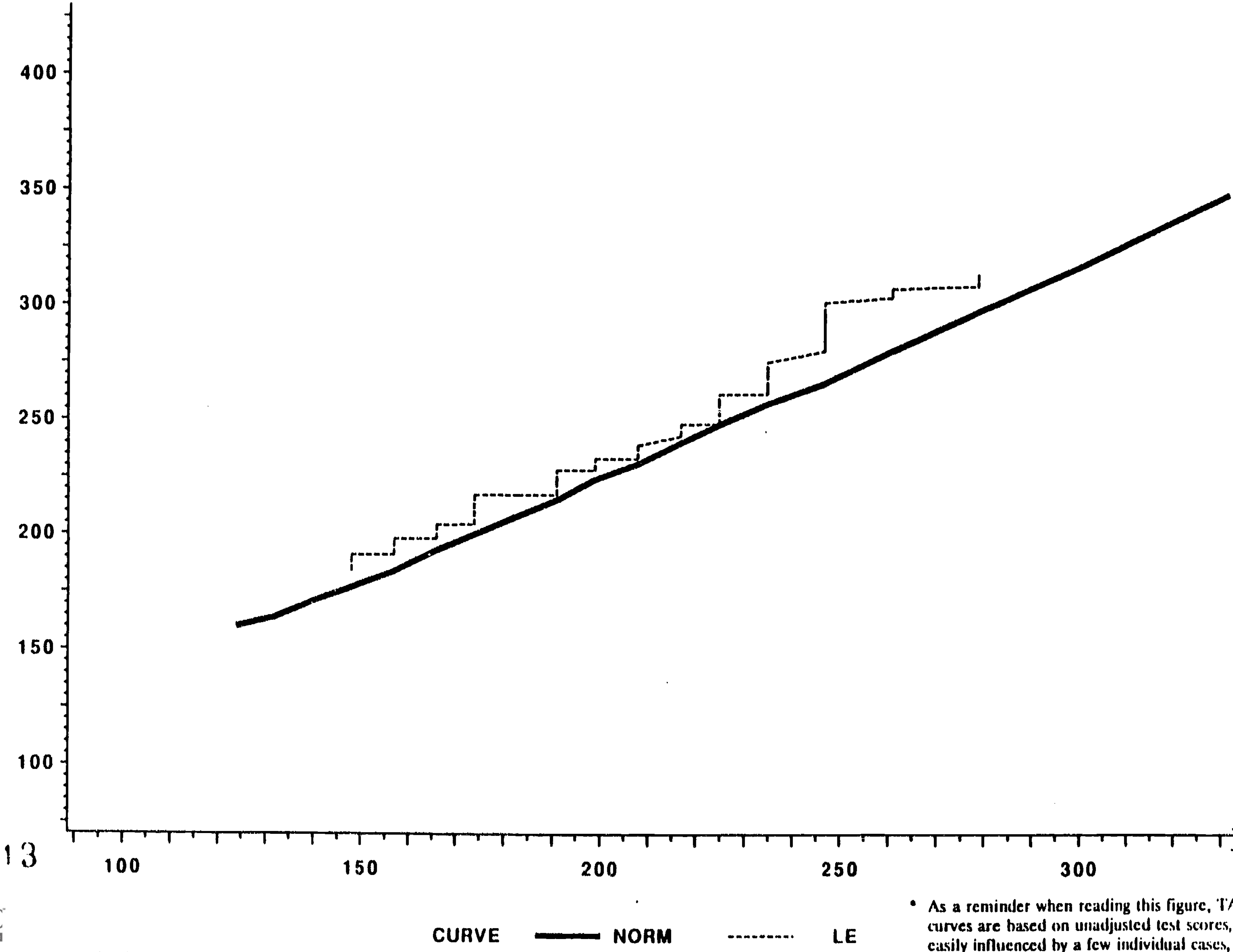


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 110

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program

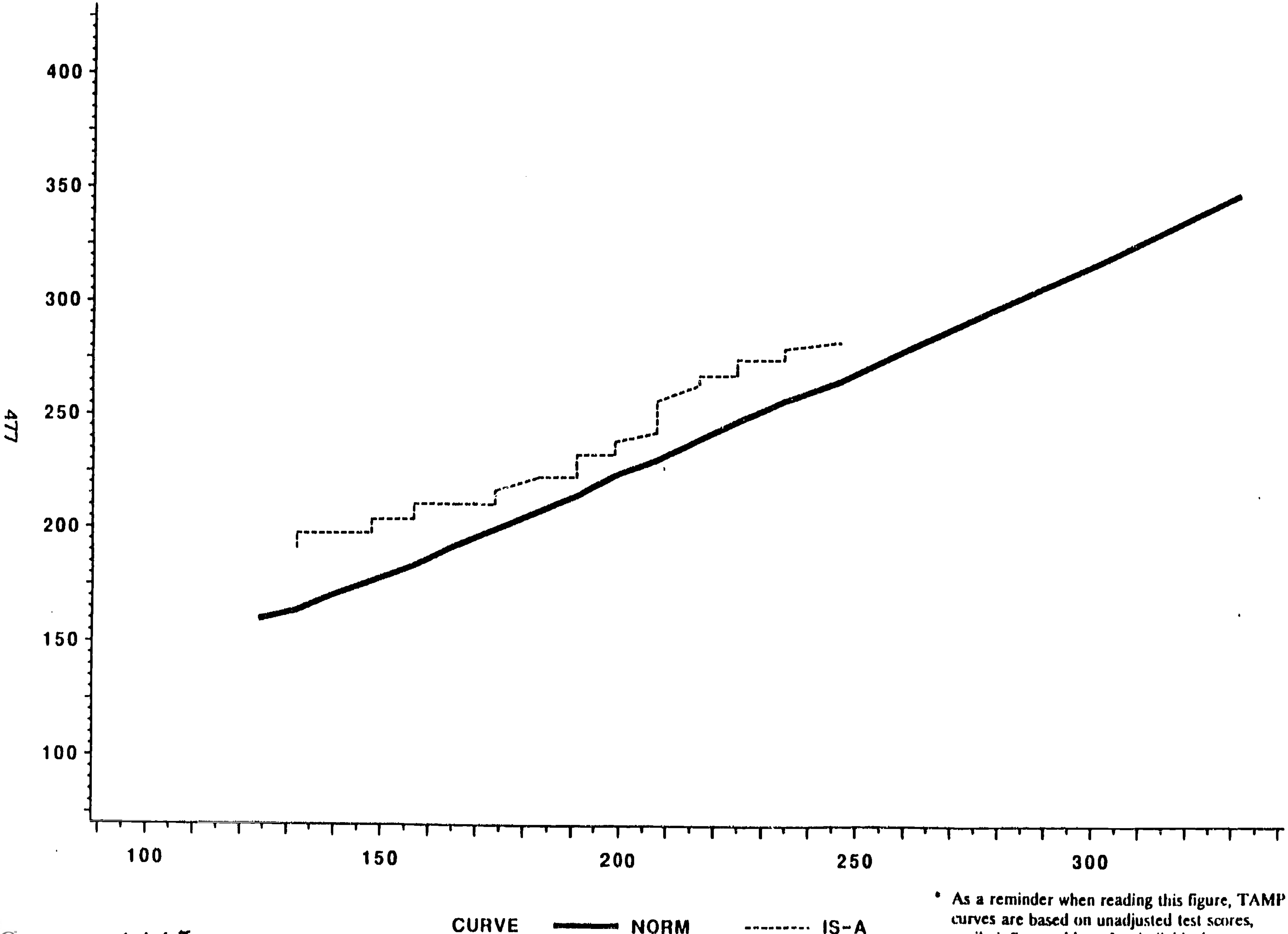


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

Figure 111

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A

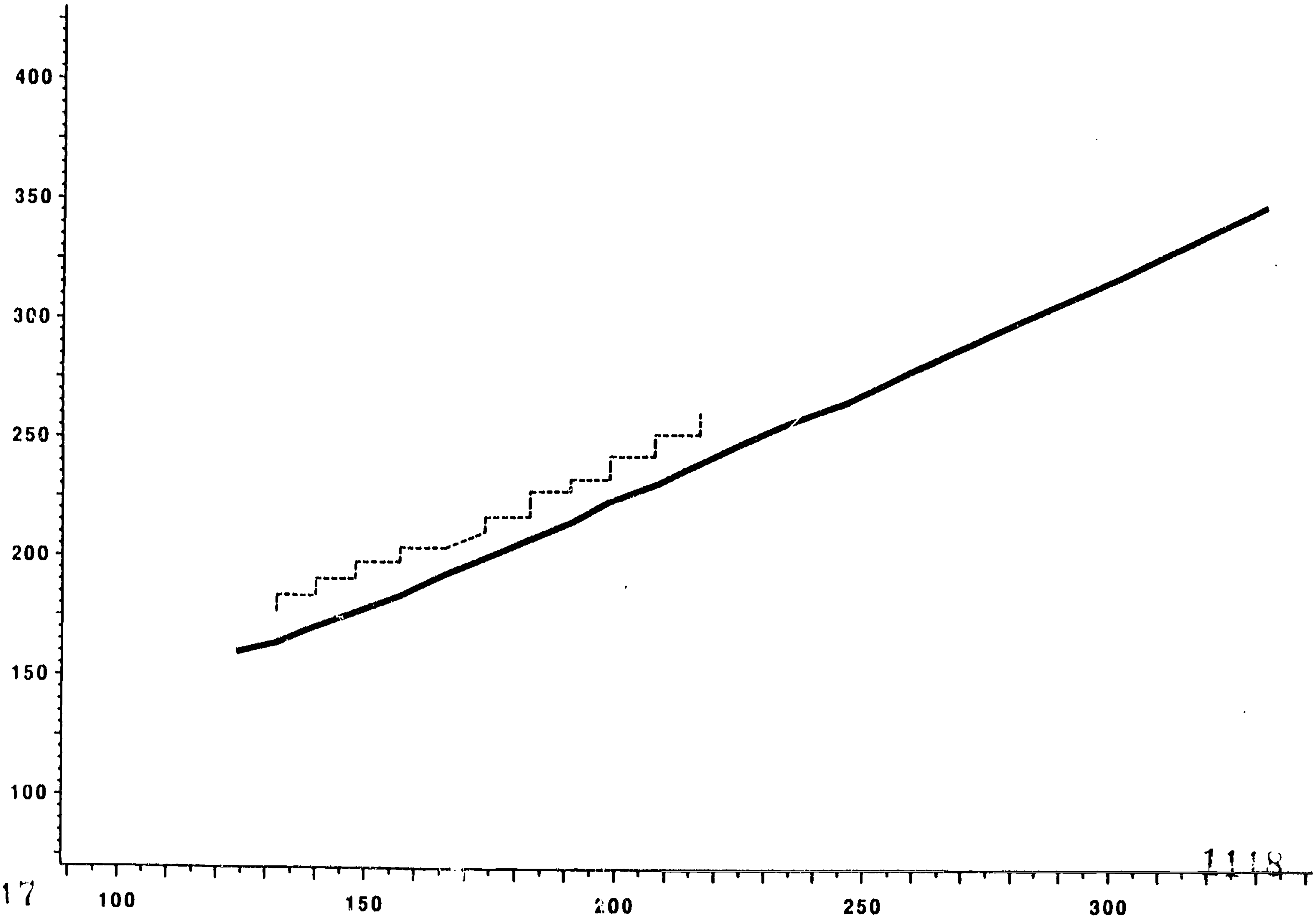


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 112

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



478

1117

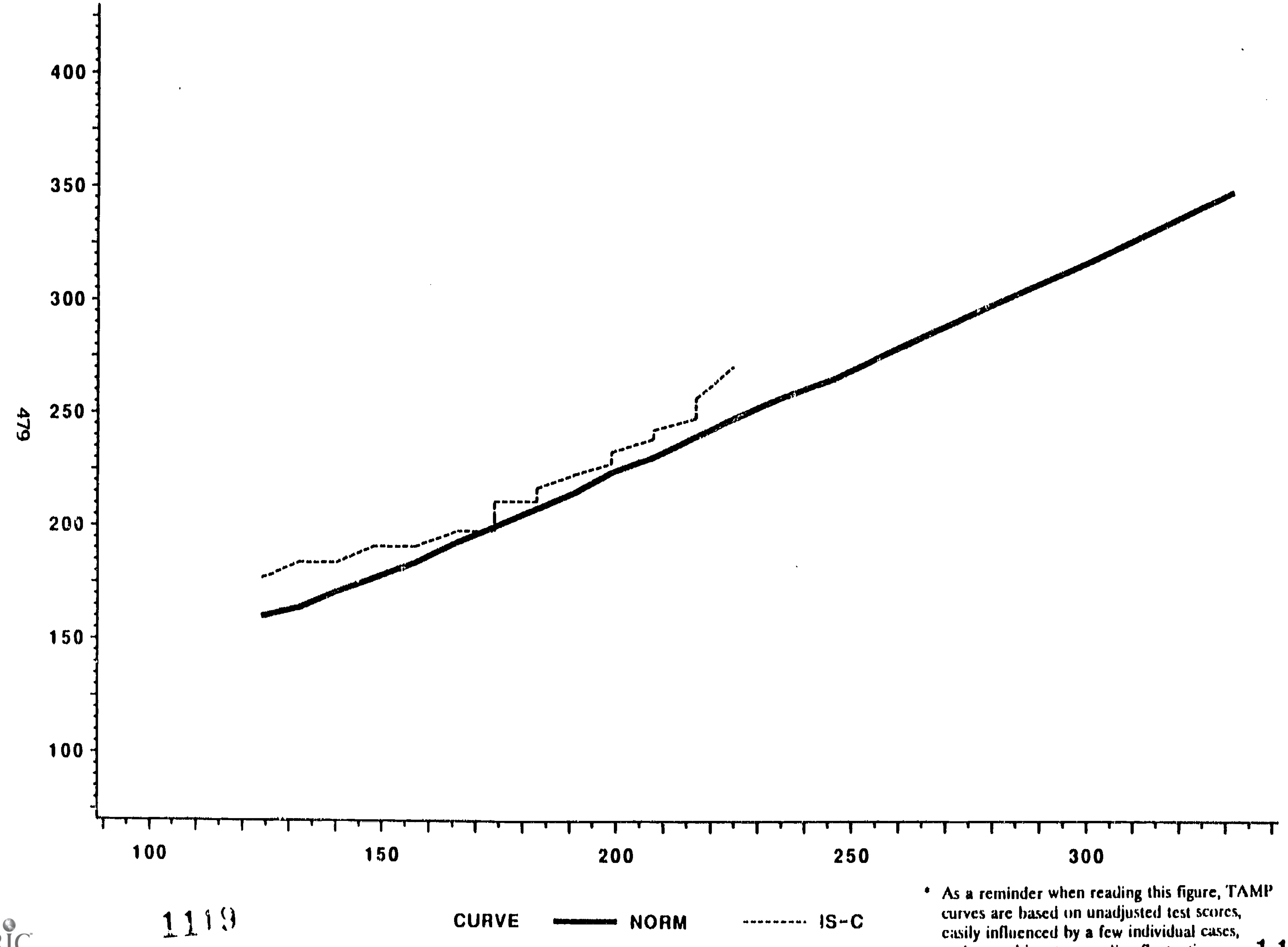
1118

• As a reminder when reading this figure, TAMP

Figure 113

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C

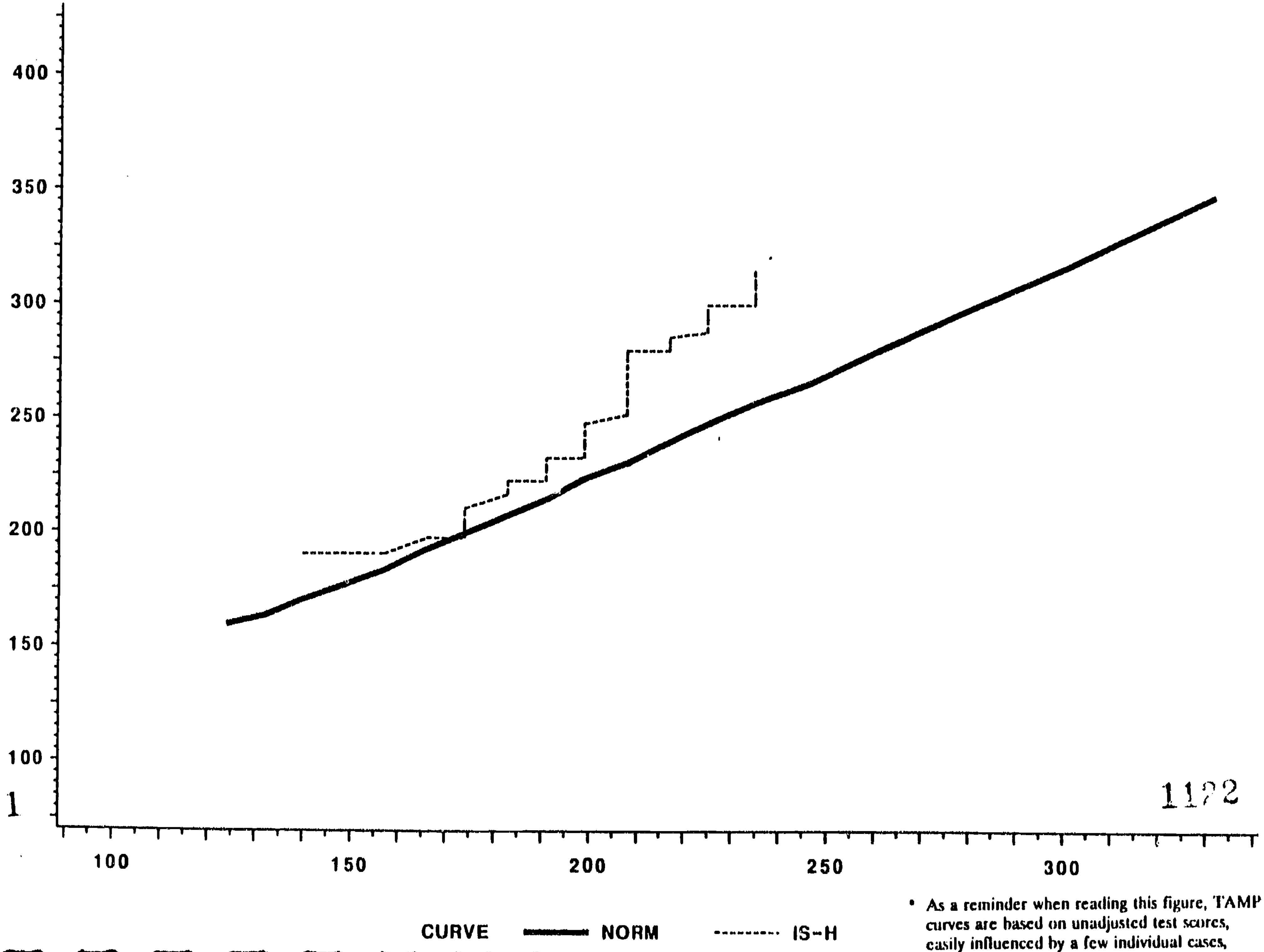


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 114

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District II

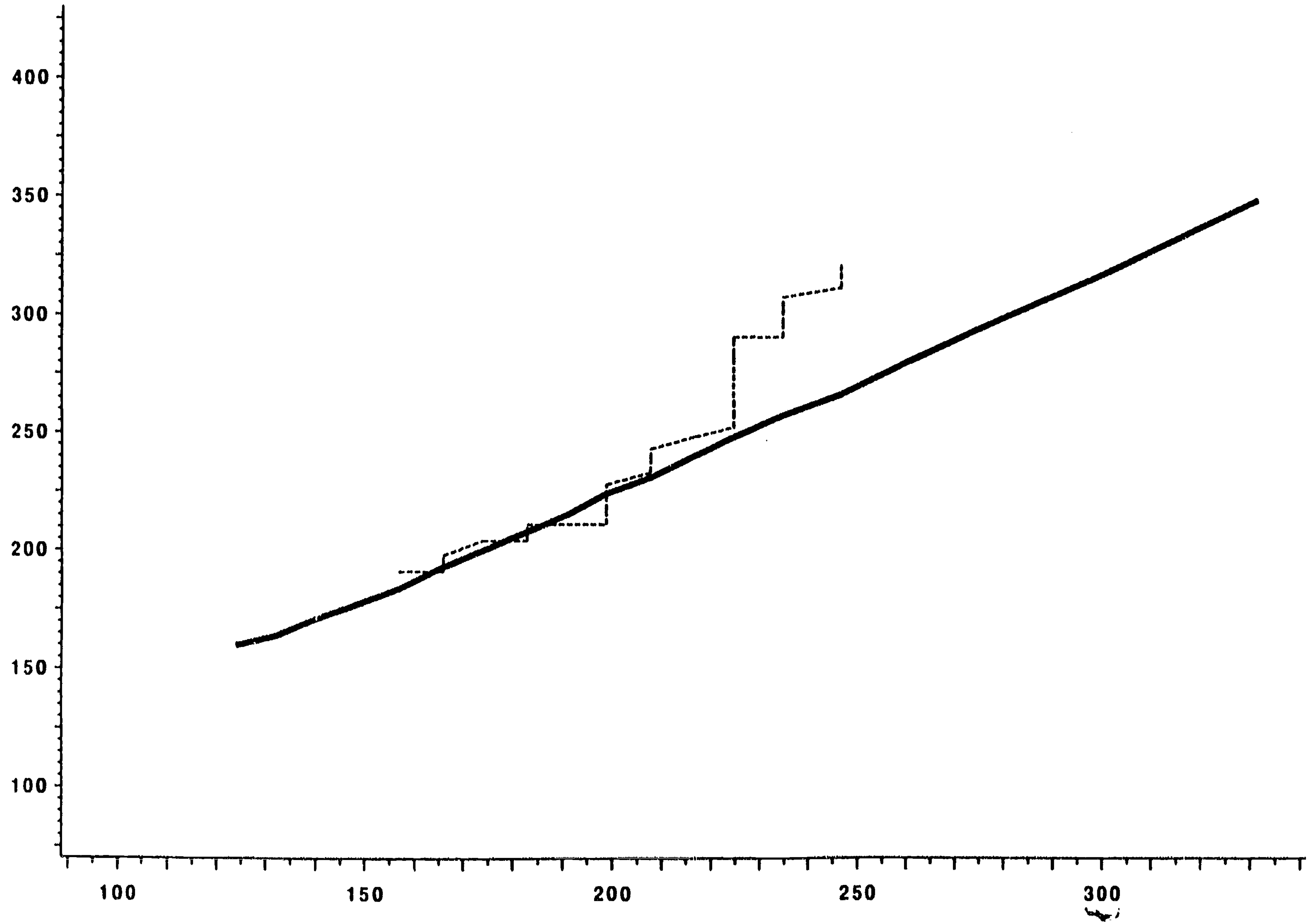


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, [redacted]

Figure 115

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A

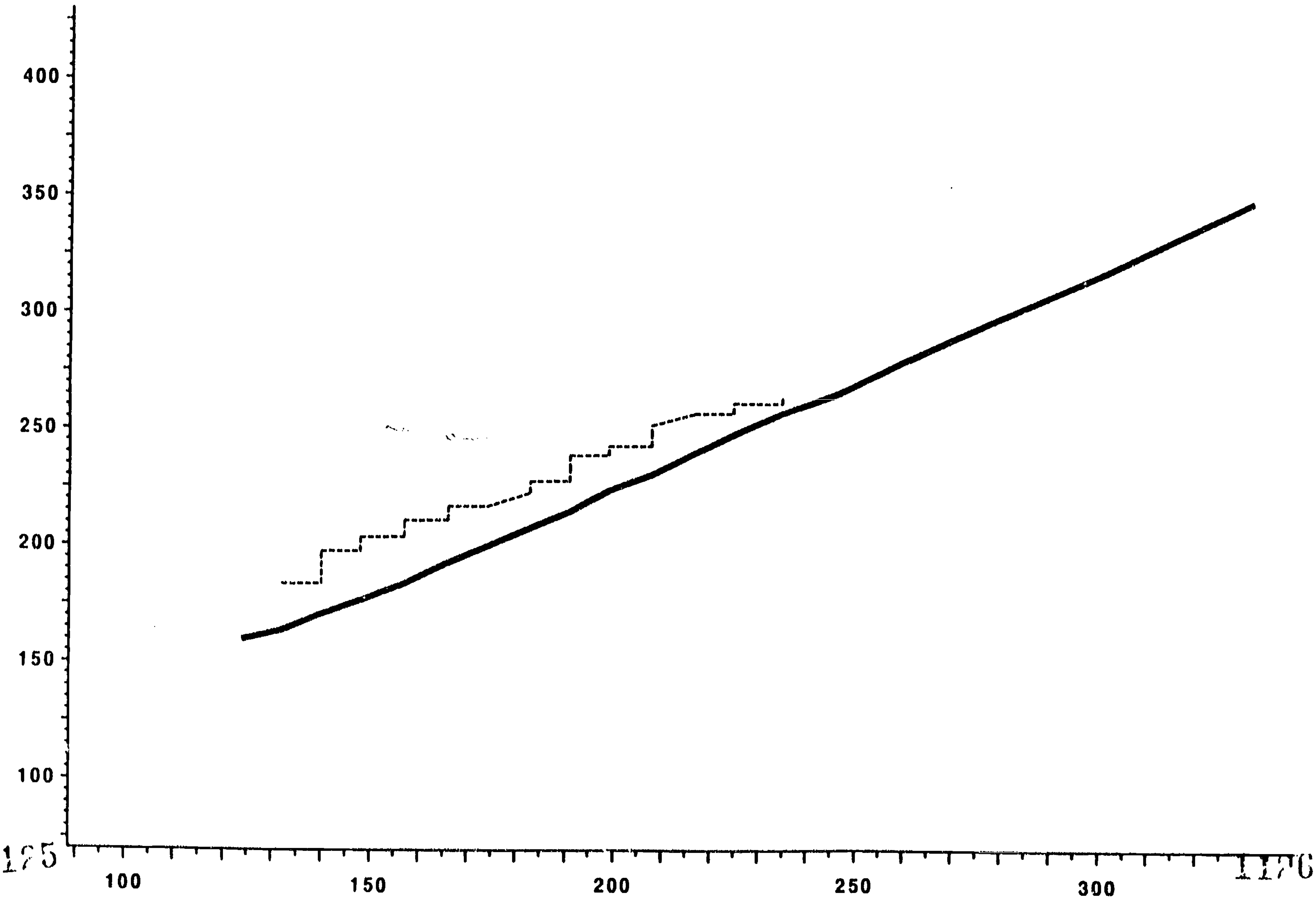


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 116

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B

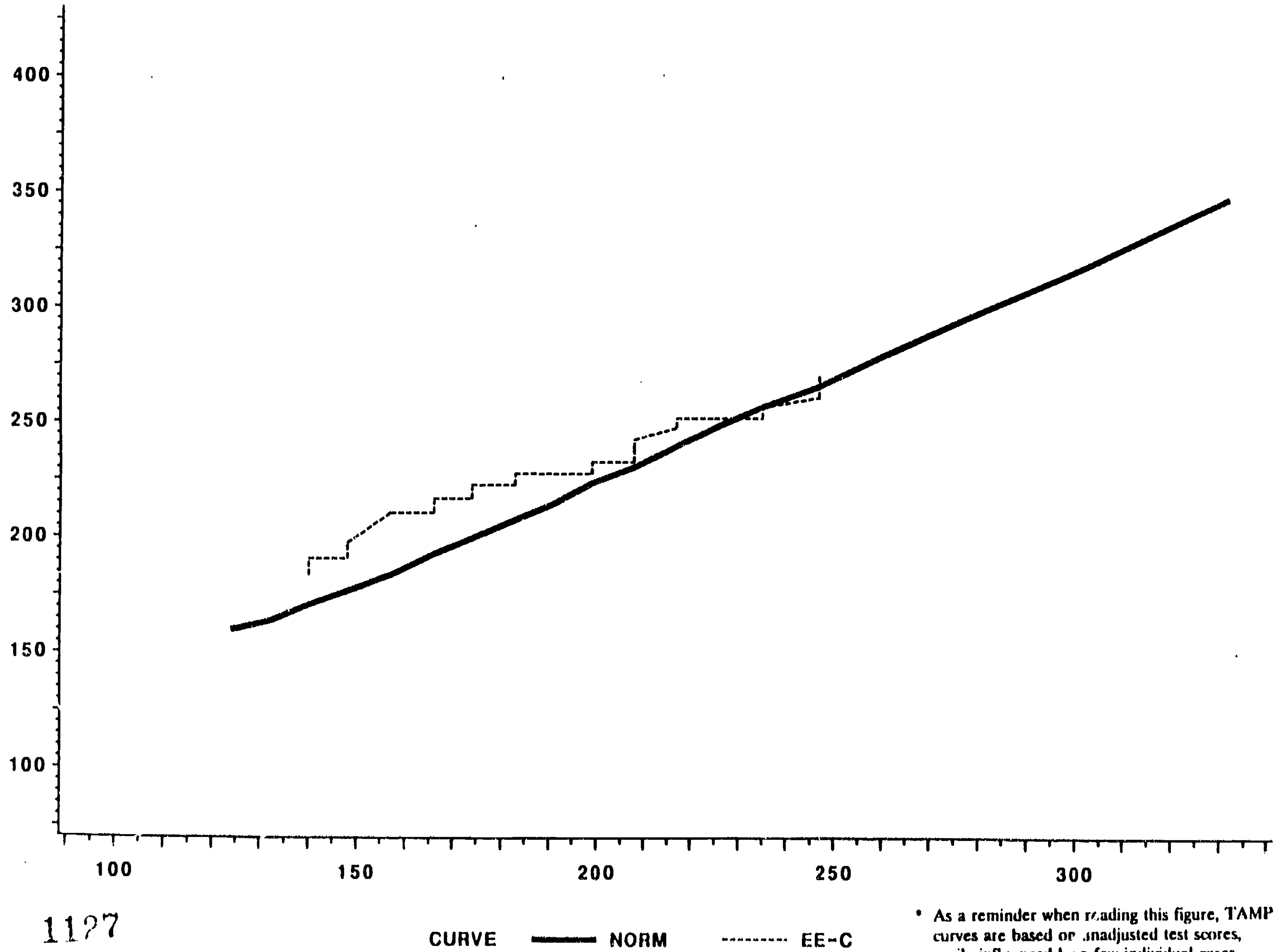


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores.

Figure 117

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C

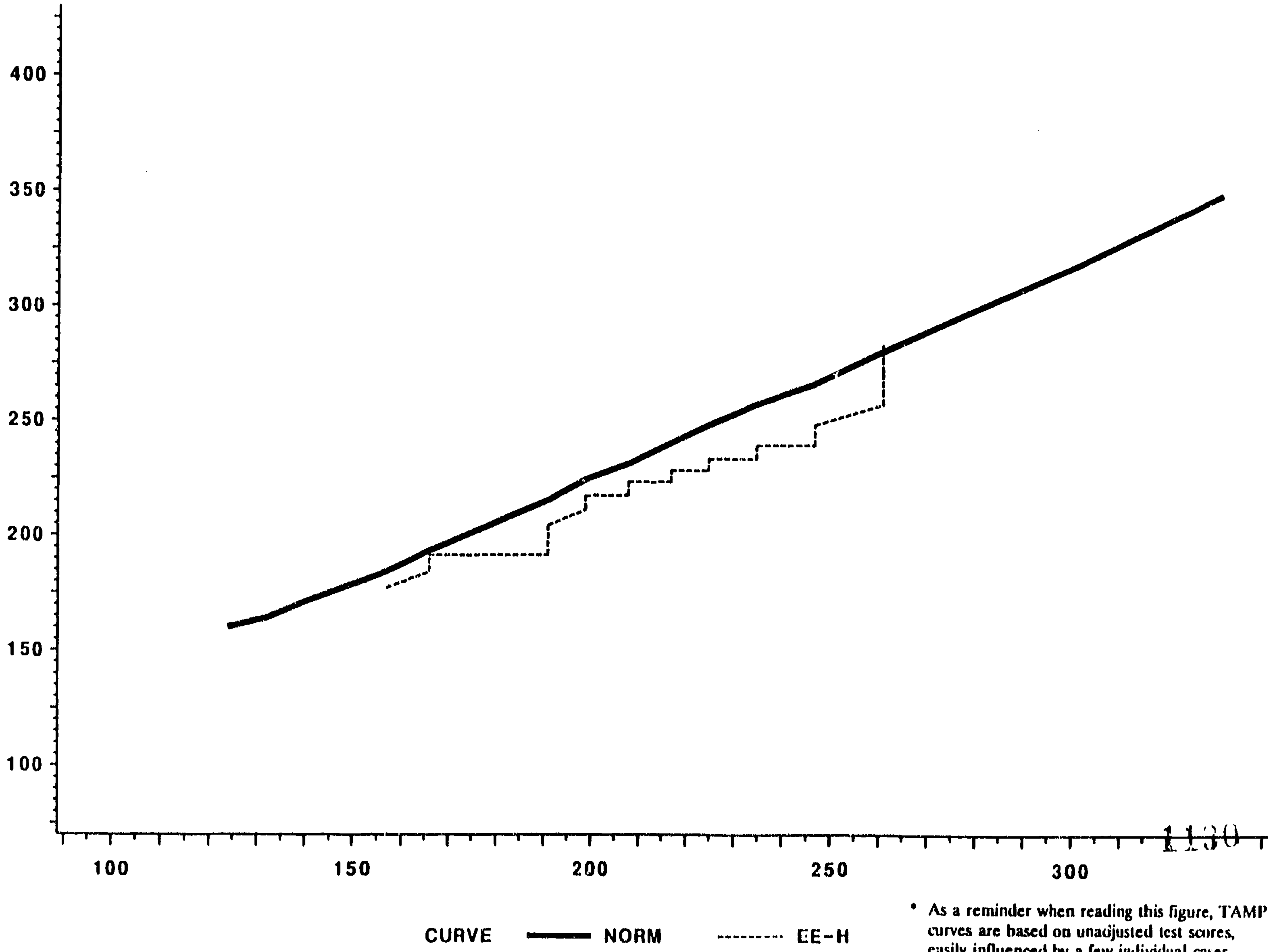


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 118

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

F.	Grade Span:	Kindergarten	to	First Grade
	Test Date:	Spring	to	Spring
	Language:	English	to	English
	Content:	Language	to	Reading

Figures 119, 120, and 121 illustrate that the growth in English reading skills seems to be comparable across the three instructional programs relative to this norming population; students in all programs appeared to be learning to read in English at a faster rate than this norming population. This suggests that if this growth rate were sustained over time all three groups would equal or surpass the average achievement of this norming population.

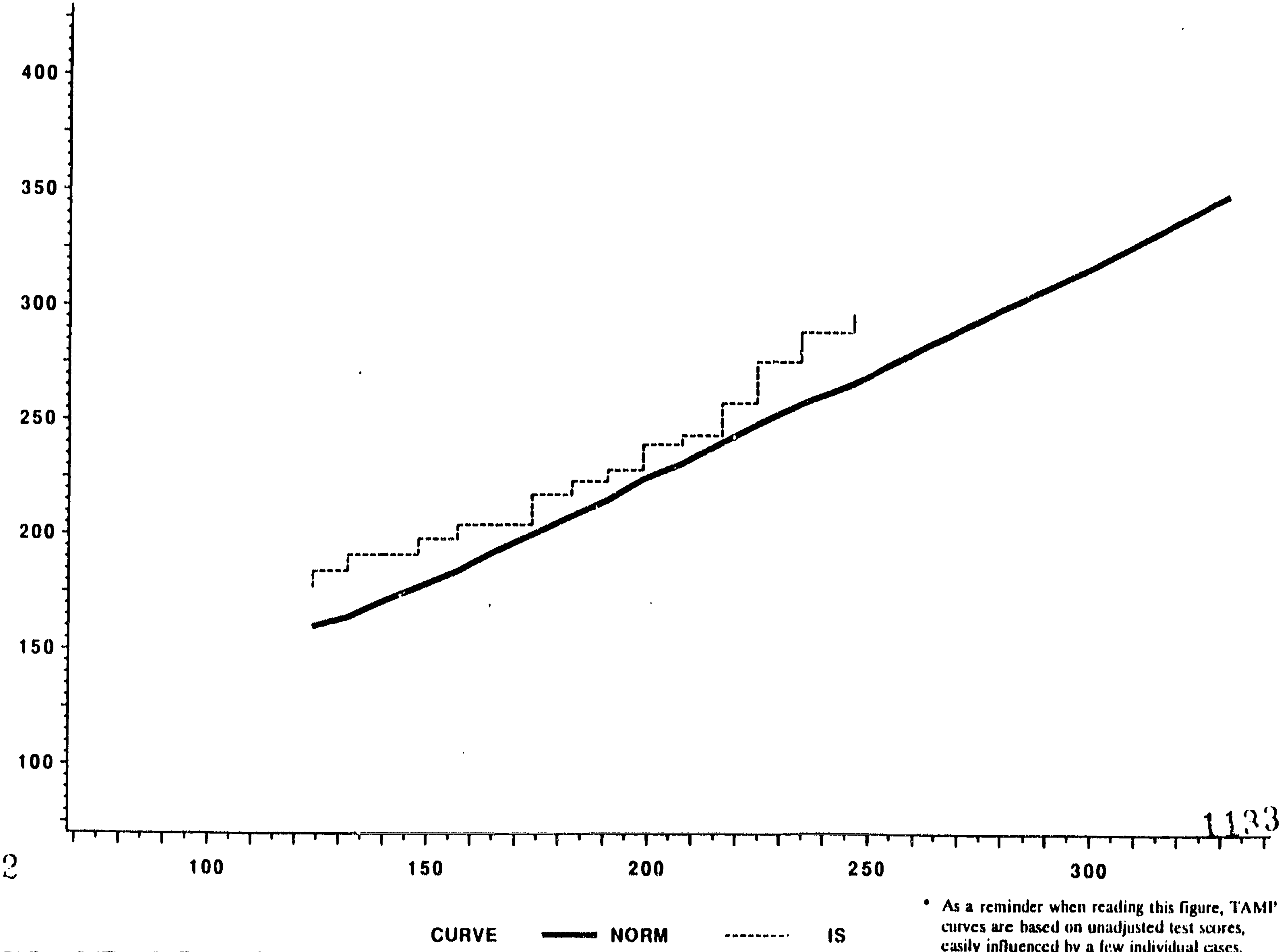
Noteworthy is that while the range of achievement for each group of students was slightly lower than for this norming population, their greater growth would seem to substantiate the value of the services for this segment of the population. Traditionally, the students consistently tend to perform lower in mainstream classes than the norming population, and they continue to lose ground relative to the norming population as they progress through school.

Little variation in growth rates was noted between the individual immersion strategy sites (see Figures 122 to 125). While no differences were noted in growth rates in early-exit sites, clearly site EE-A had a higher distribution of spring kindergarten scores than the other early-exit sites (see Figures 126 to 129). Finally, it appears that site EE-H students again grew the slowest relative to this norming population.

Figure 119

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, nd a [illegible] ject [illegible] plin [illegible] uati [illegible]

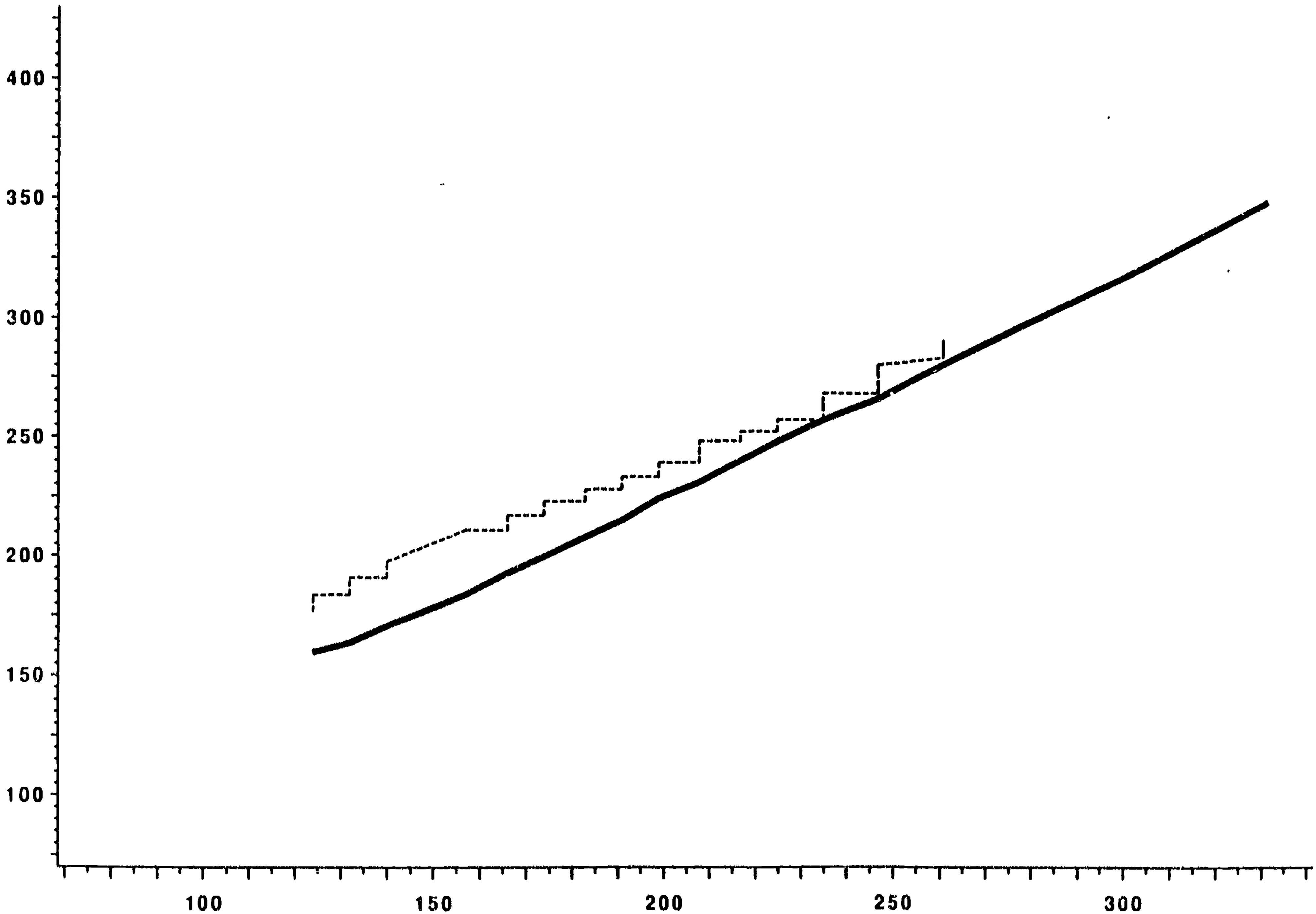
1132 1133

ERIC

Figure 120

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



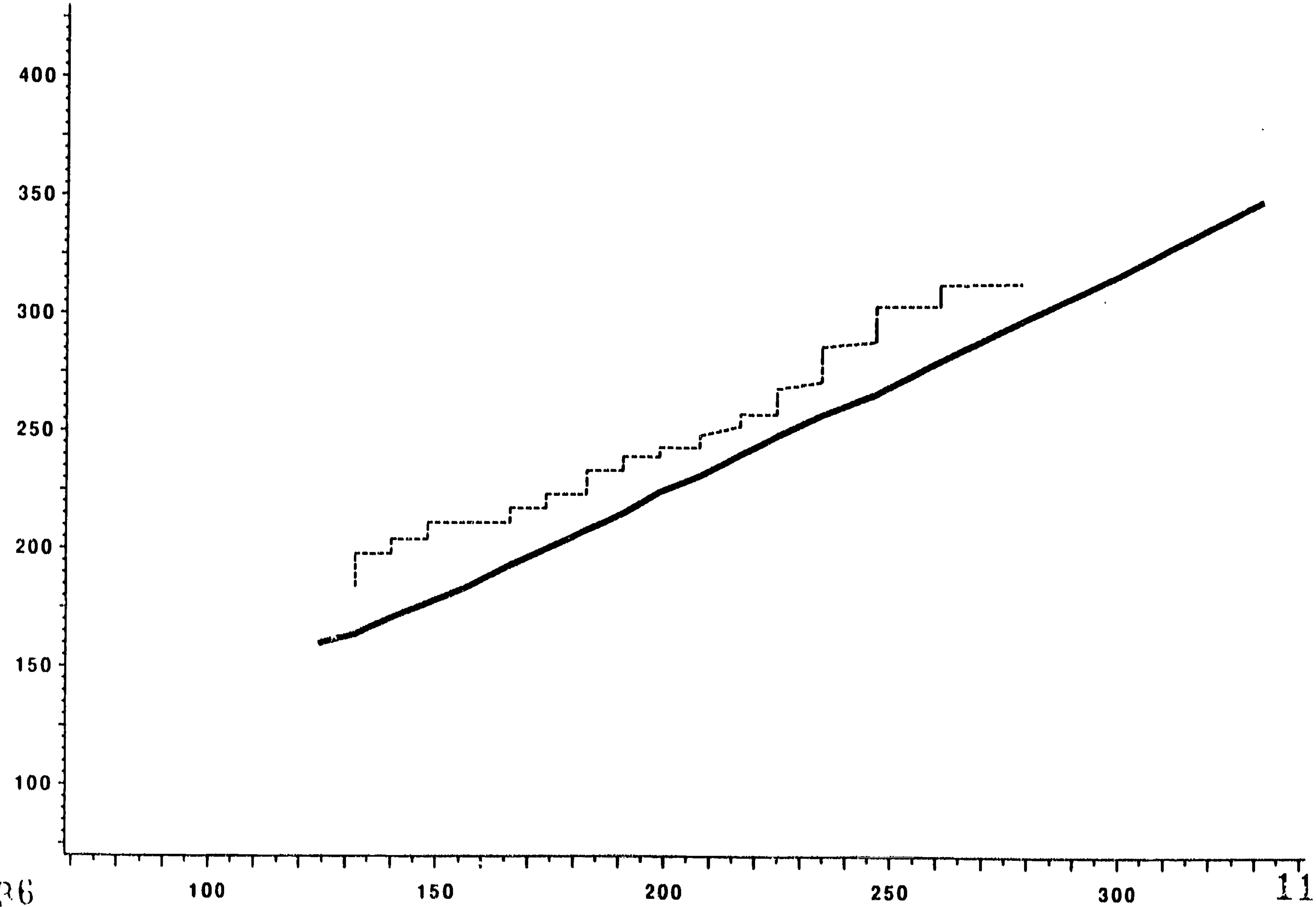
CURVE NORM EE

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

Figure 121

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program

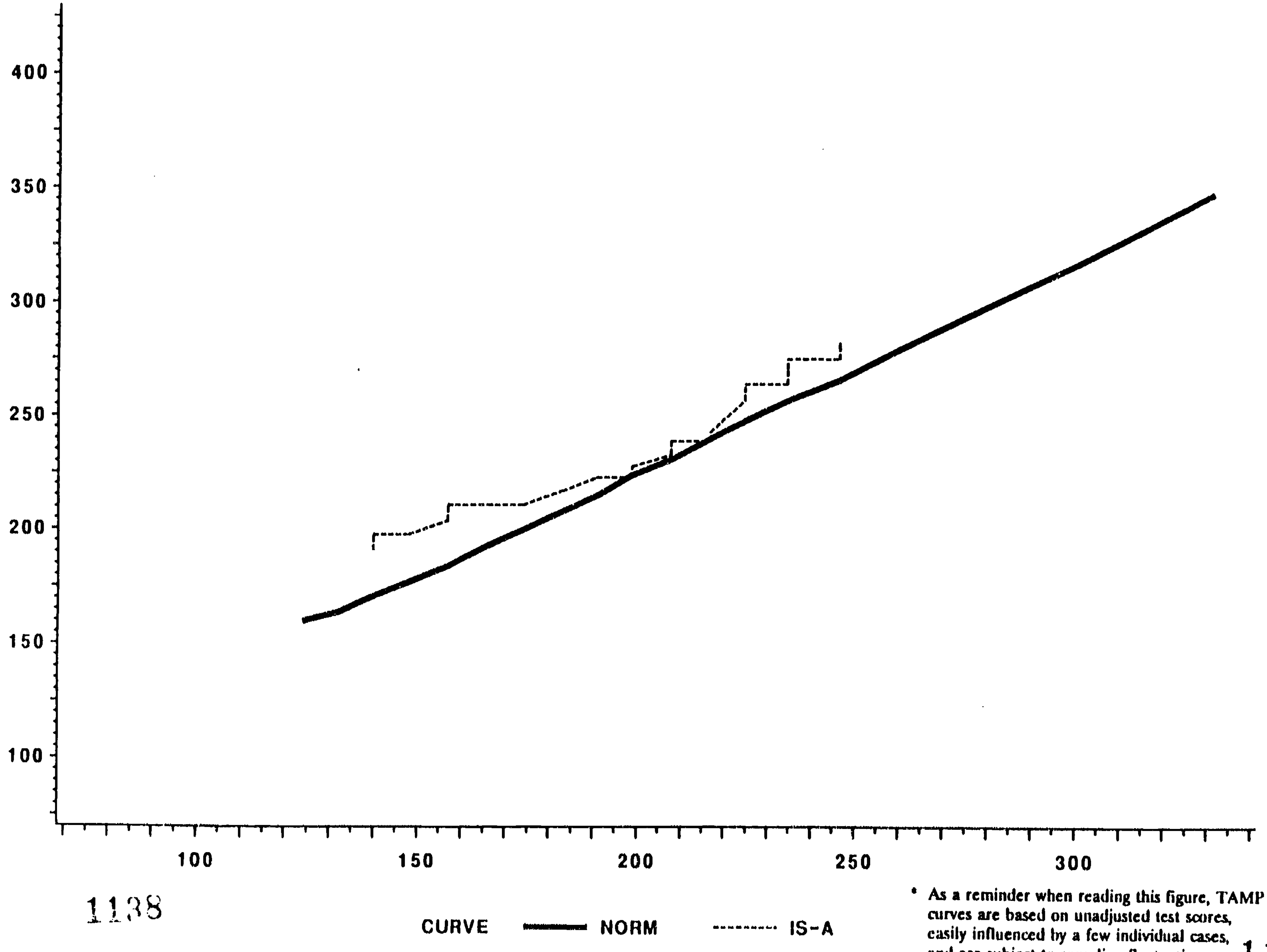


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

Figure 122

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



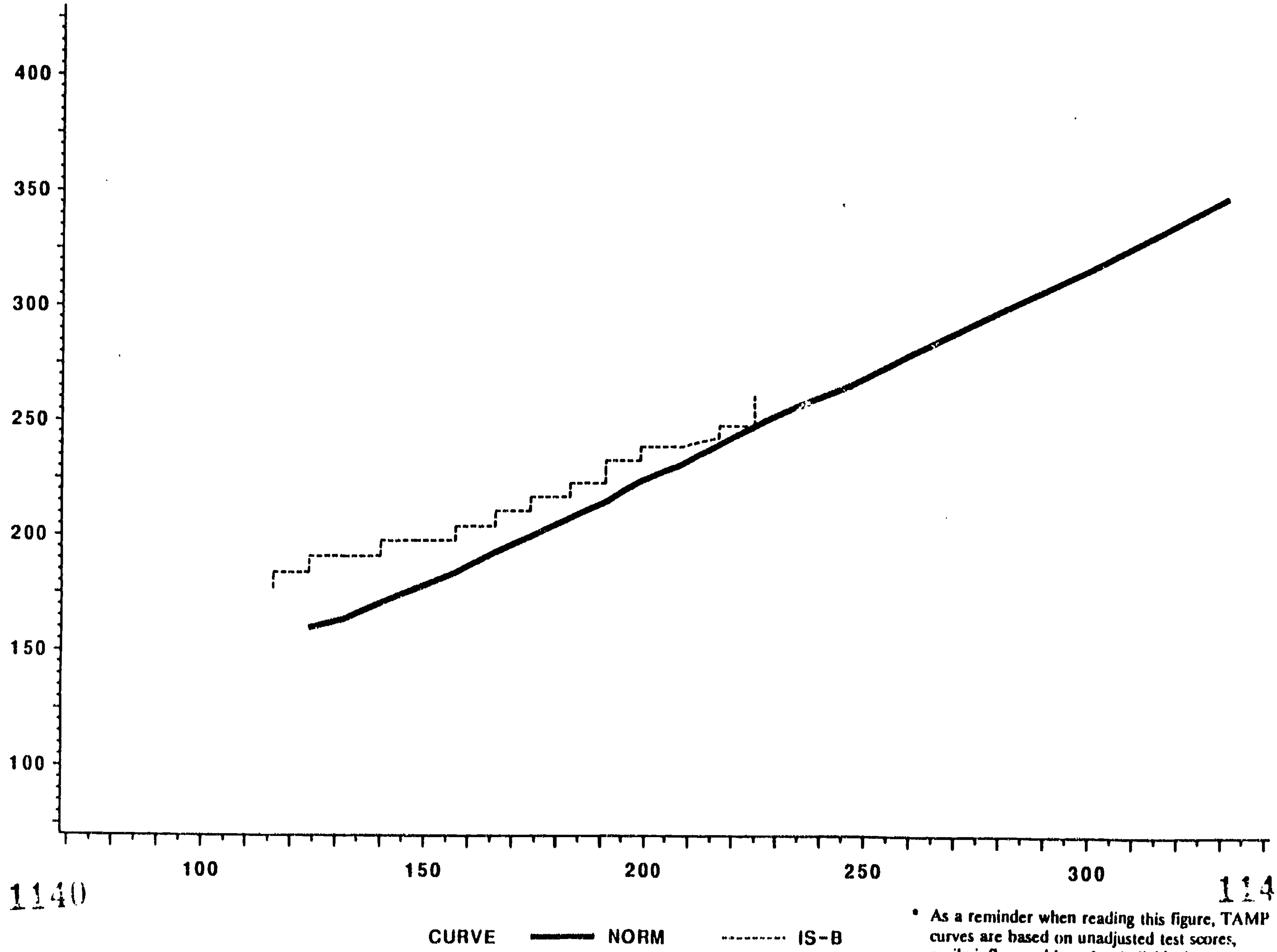
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 123

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



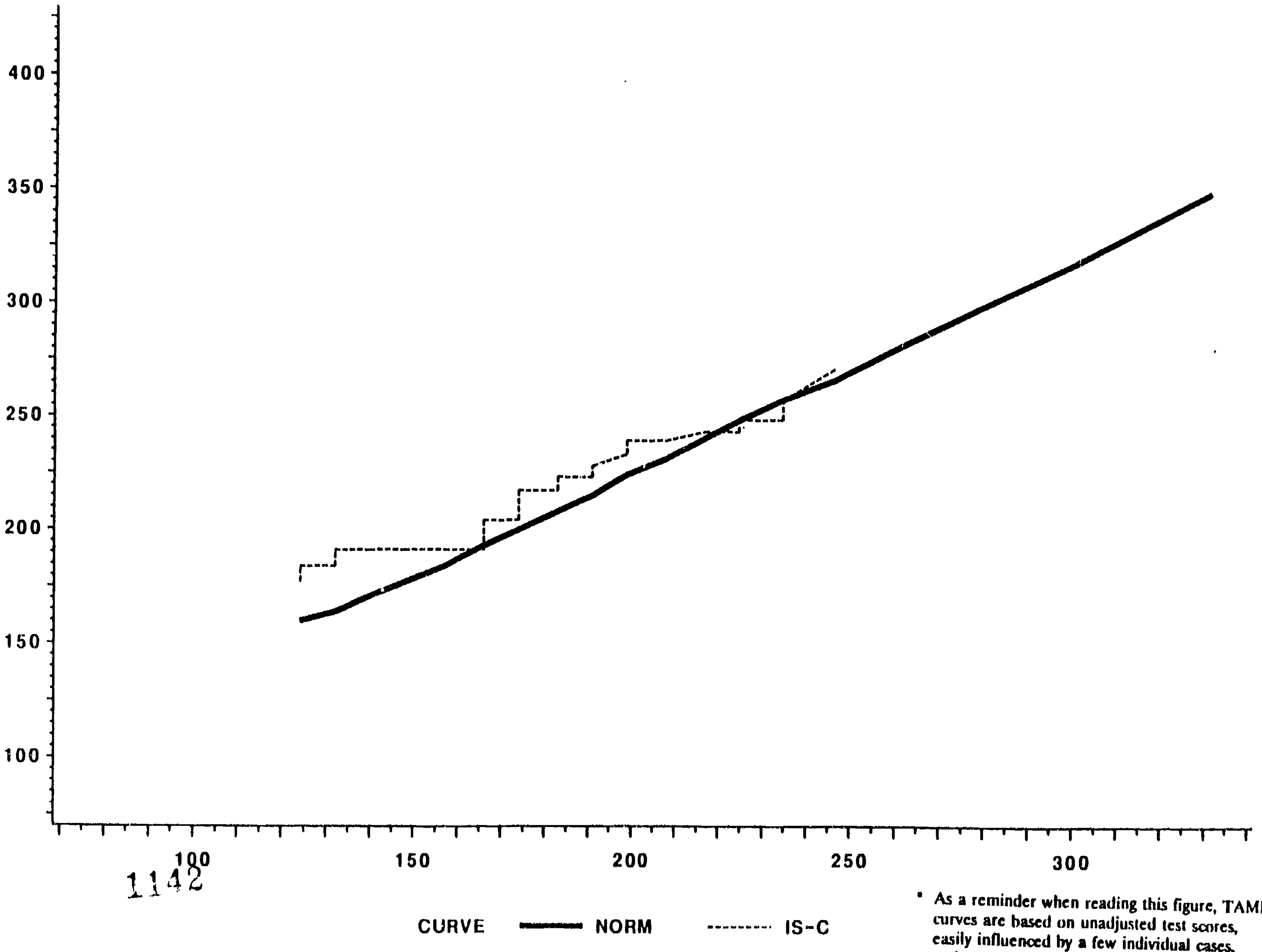
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 124

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

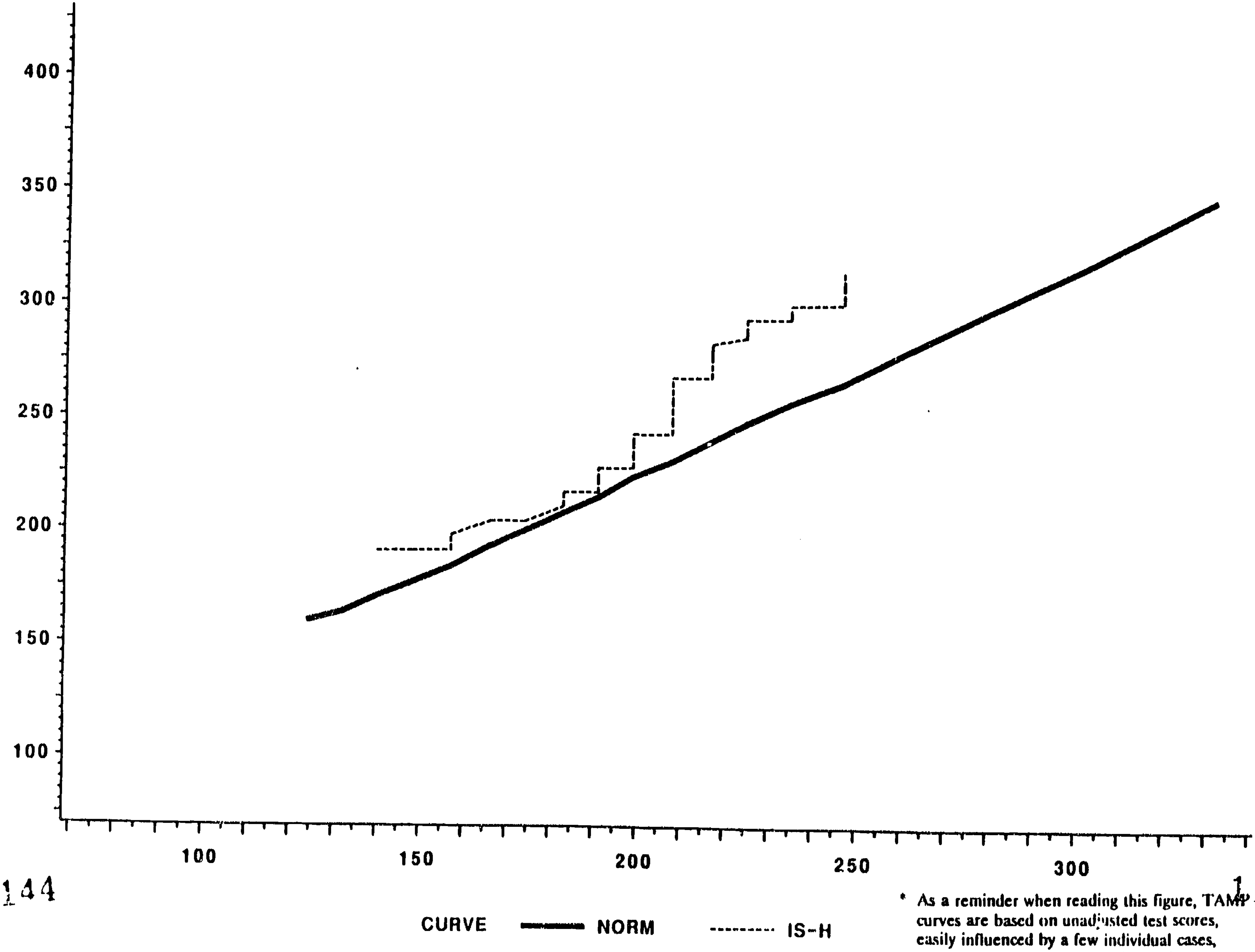
491

ERIC

Figure 125

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H

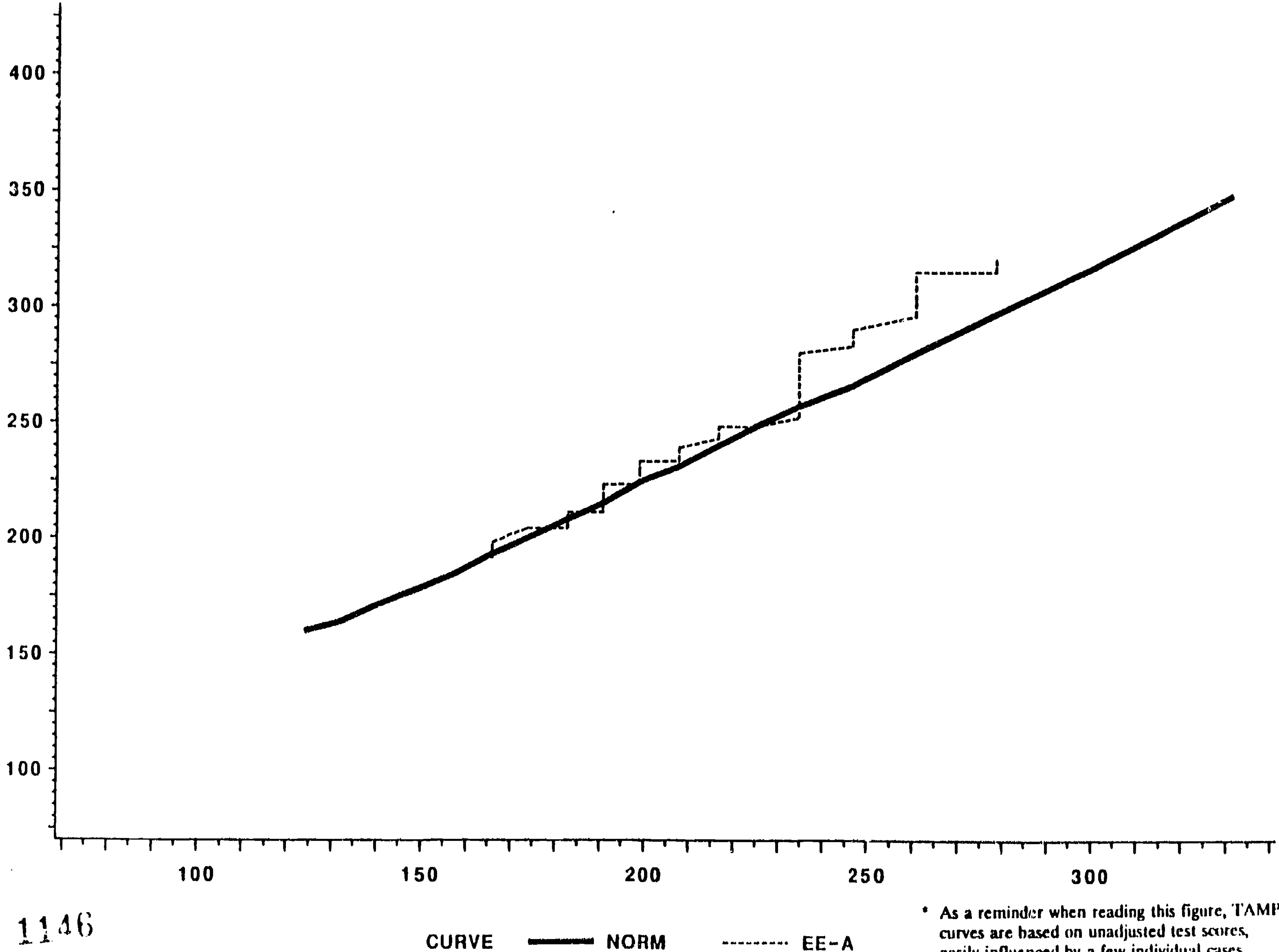


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 126

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A

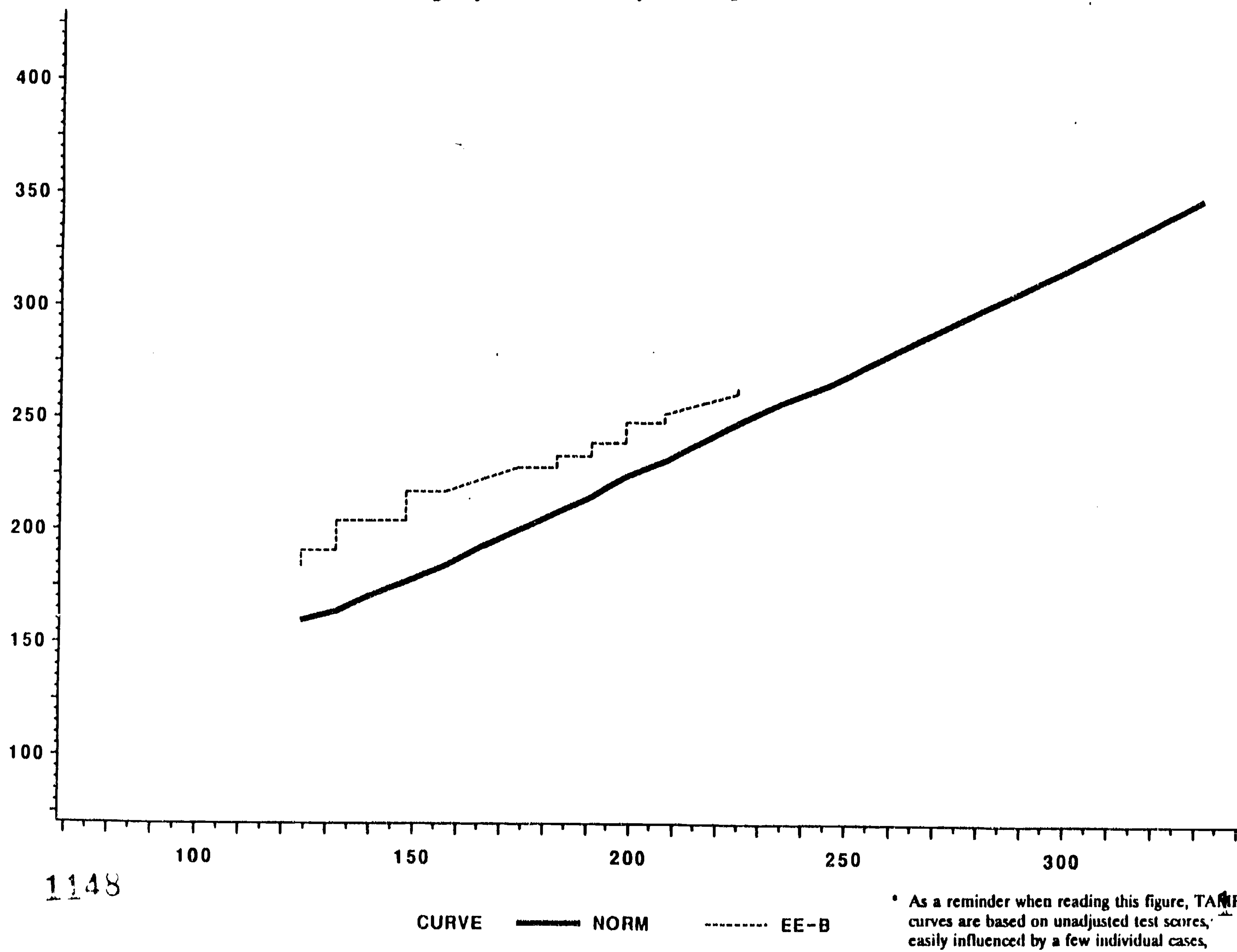


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 127

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



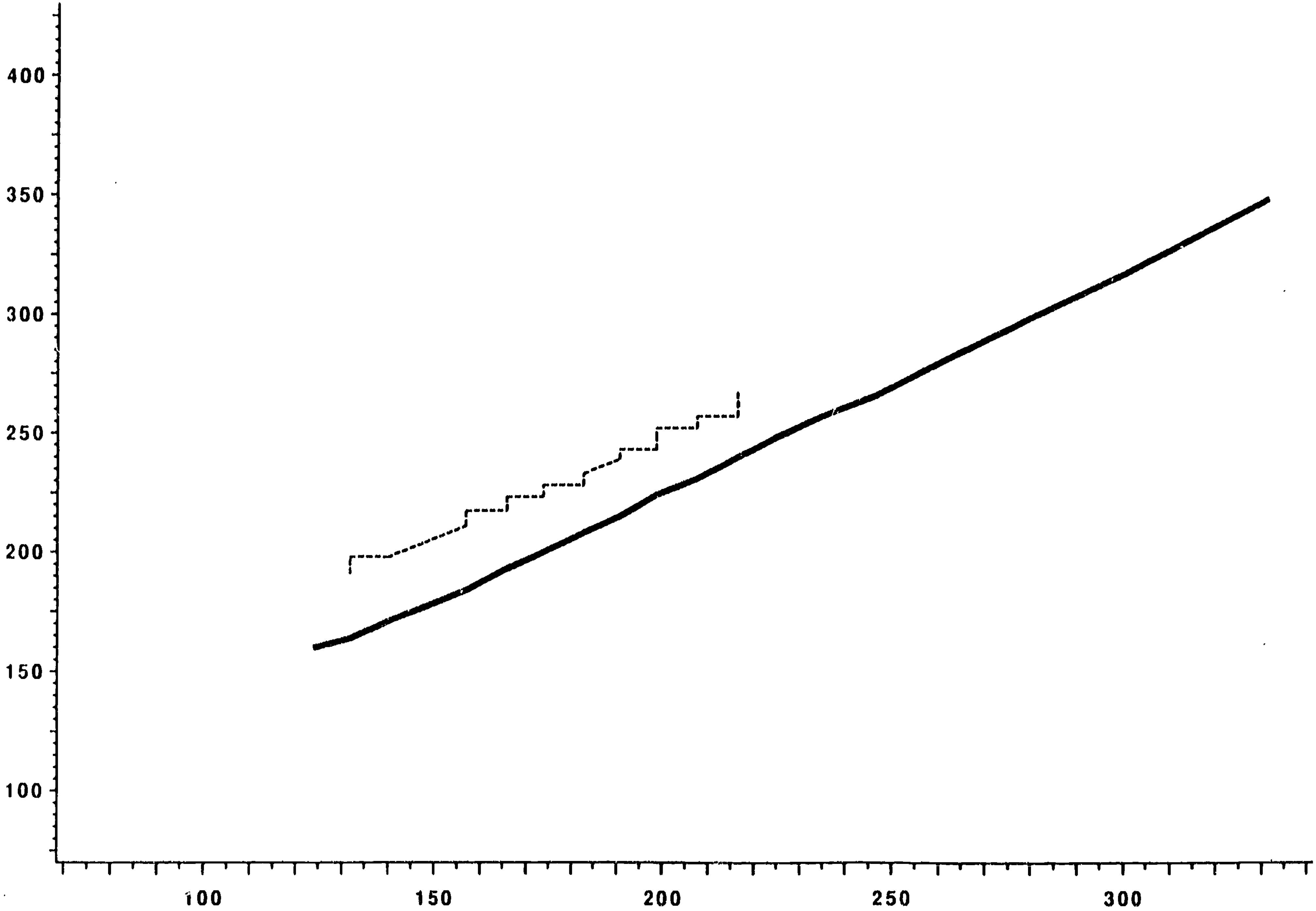
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

Figure 128

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C

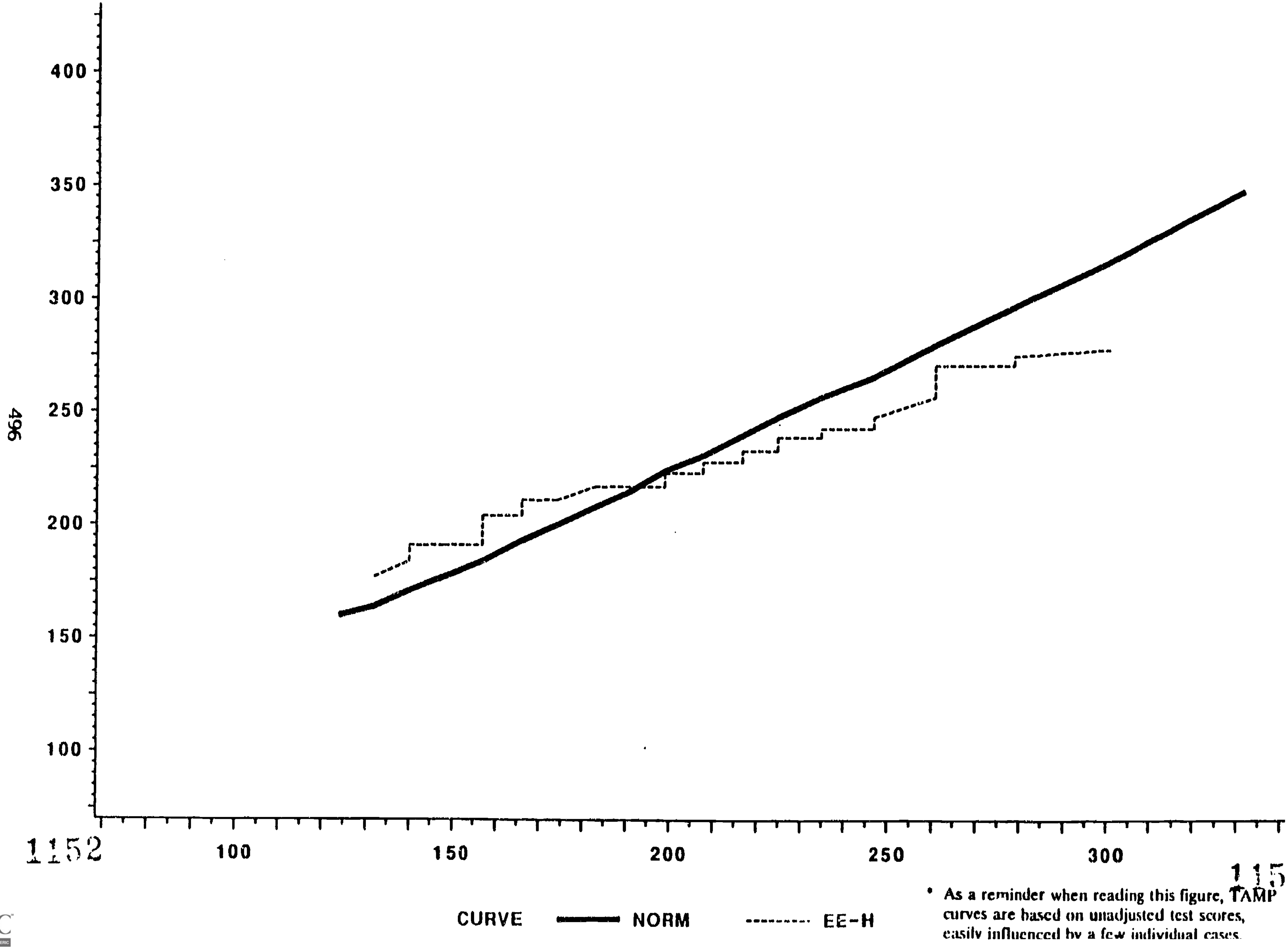


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 129

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

3. Fall Kindergarten to Spring First Grade:

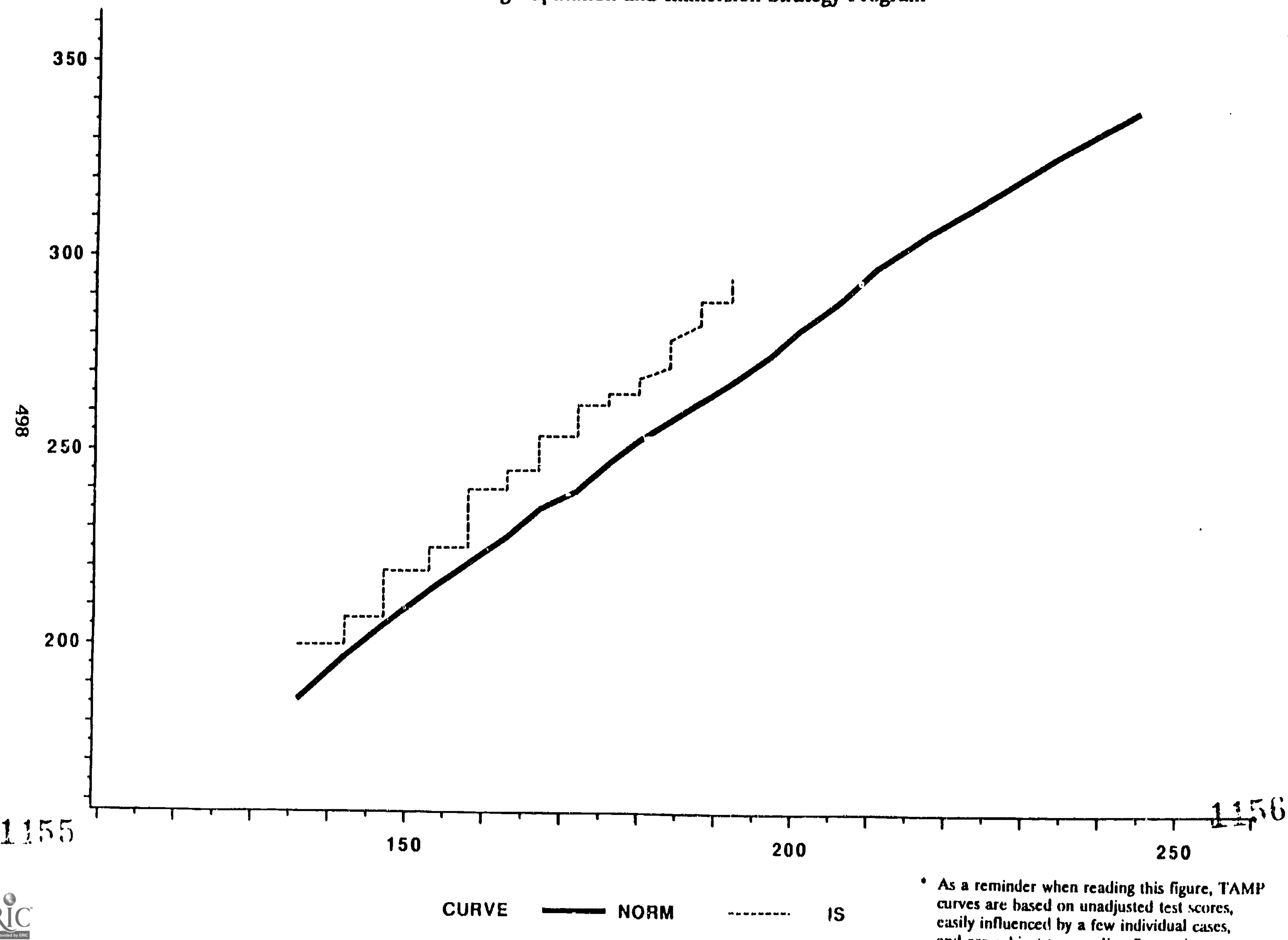
A.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

Figures 130, 131, and 132 address two issues. First, to what extent were target students' entry-level skills comparable to those of this norming population? Secondly, how comparable were the growth rates in mathematics among immersion strategy, early-exit, and late-exit target students from fall kindergarten when tested in Spanish to spring first grade when tested in English? Figures 130, 131, and 132 suggest that while students across programs had lower entry-level mathematics skills relative to this norming population, they all seemed to learn at a faster rate than comparable students in this norming population. It seems that teachers in all three programs were very successful in helping target students develop mathematics skills. Examination of the TAMP curves for the individual sites for immersion strategy (see Figures 133 to 136) and early-exit (see Figures 137 to 140) suggest little between-site variation within the IS or EE programs. The only exception is site IS-A wherein students in general did not seem to grow faster than this norming population, and students with the lowest entry-level skills seemed to grow more slowly than this norming population.

Figure 130

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



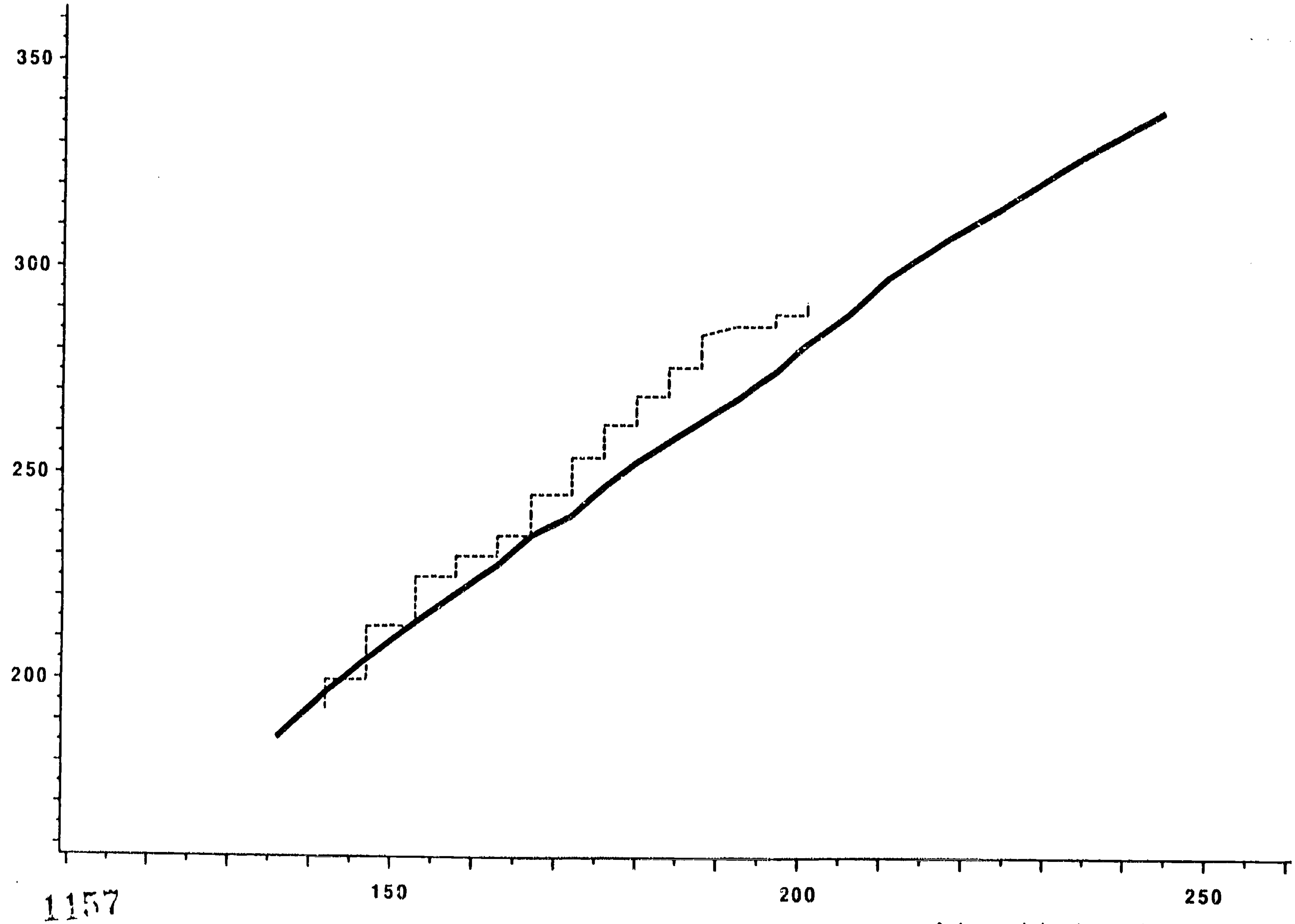
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 131

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

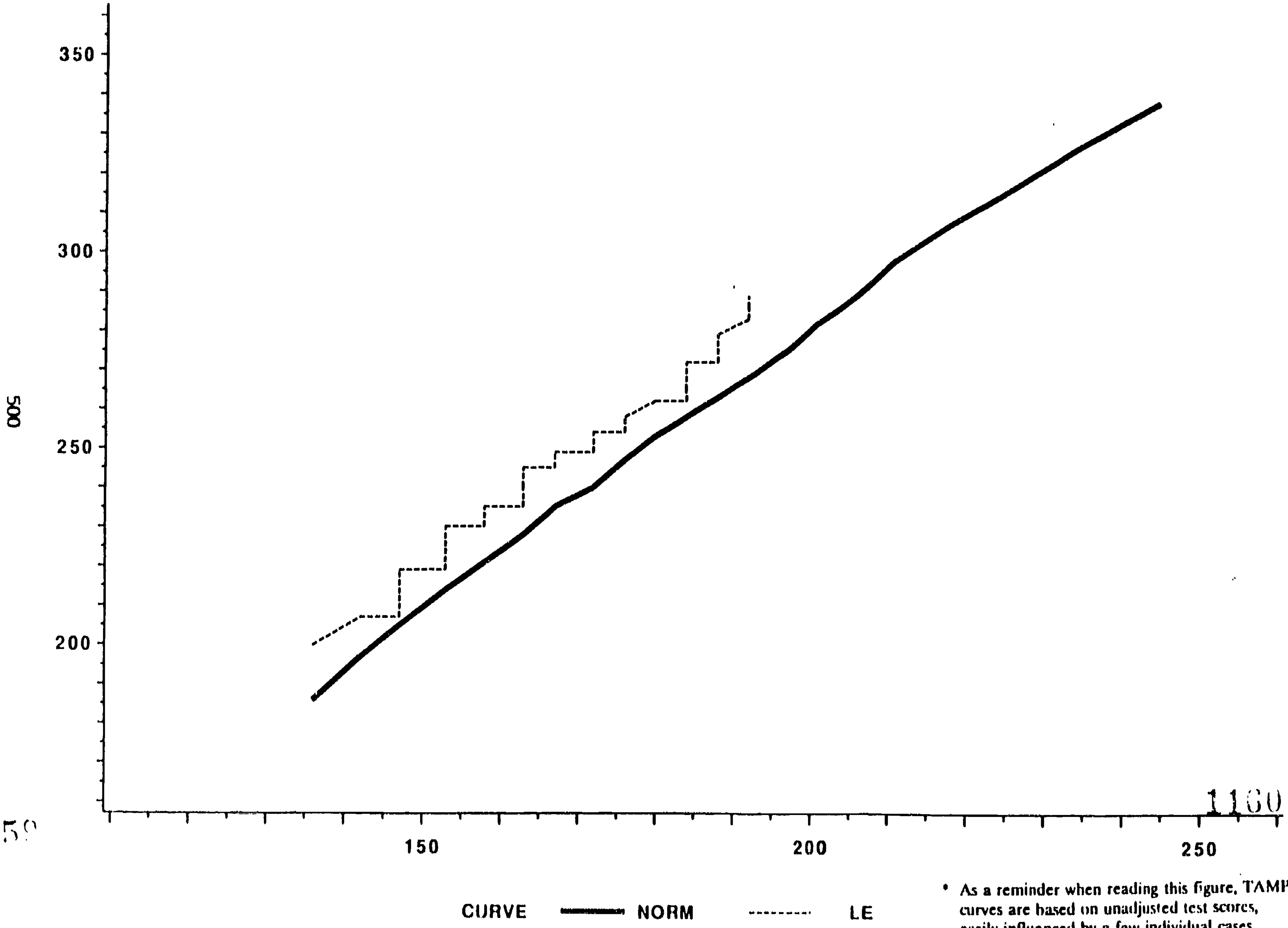


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 132

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



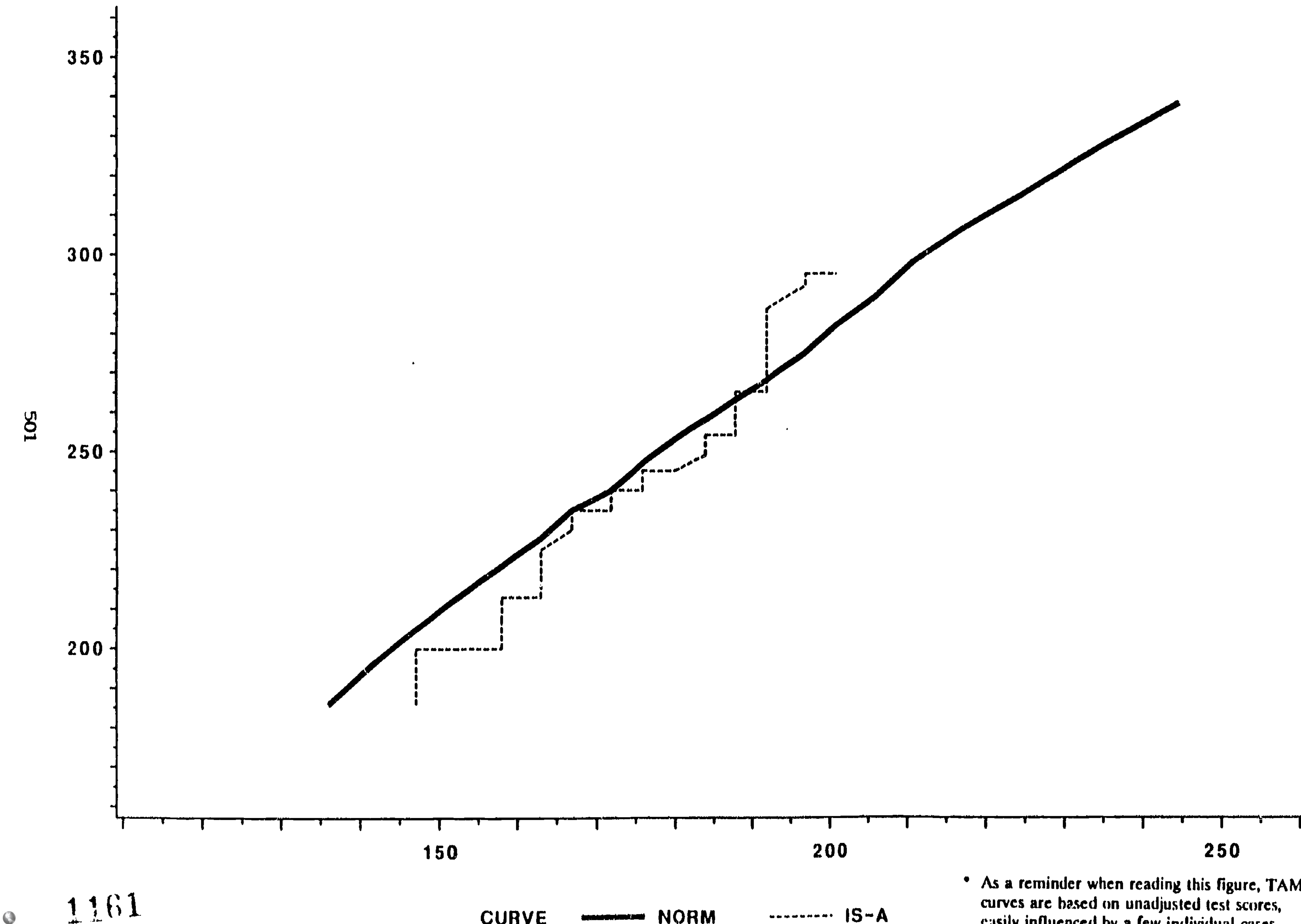
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 133

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



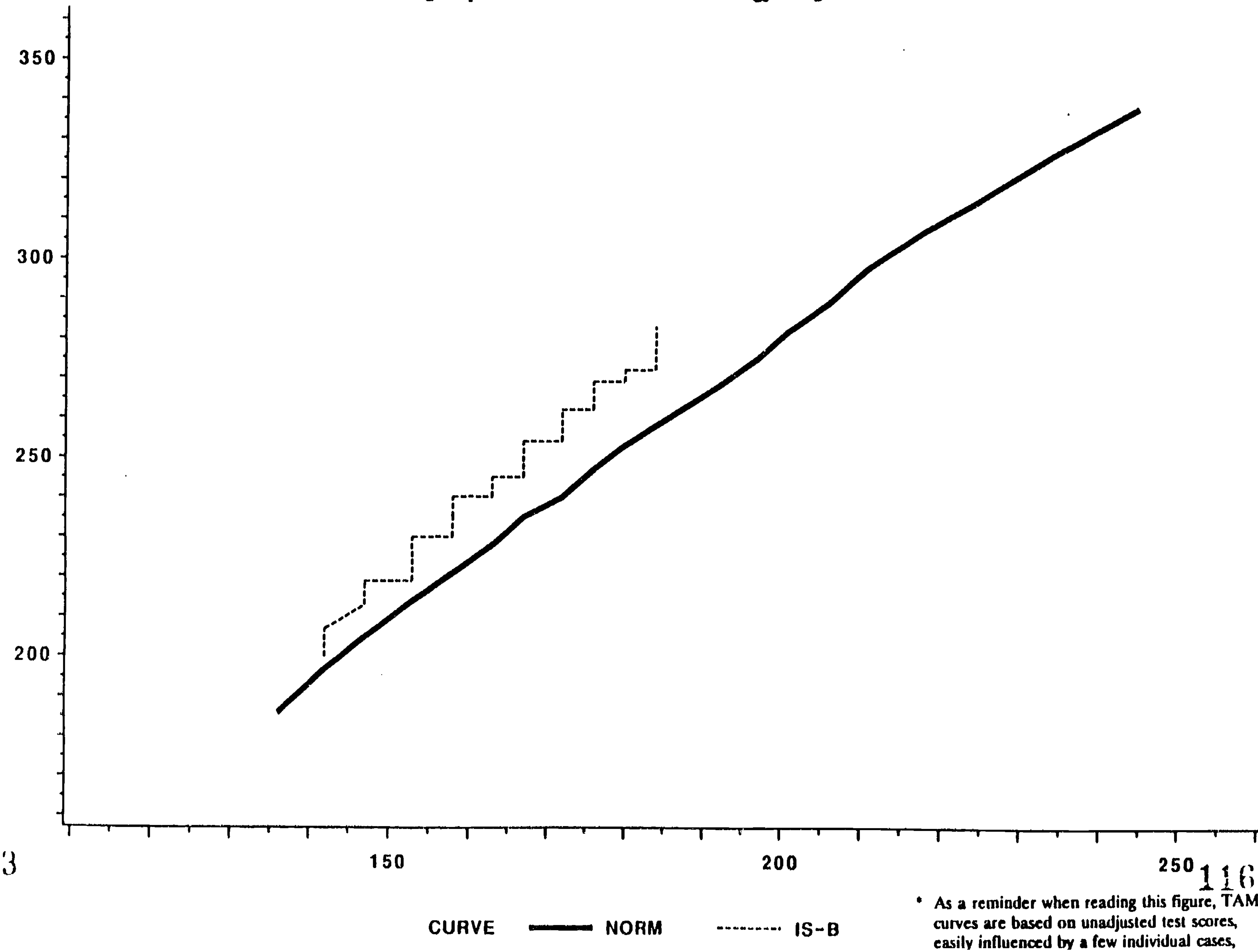
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 134

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



502

1163 1164

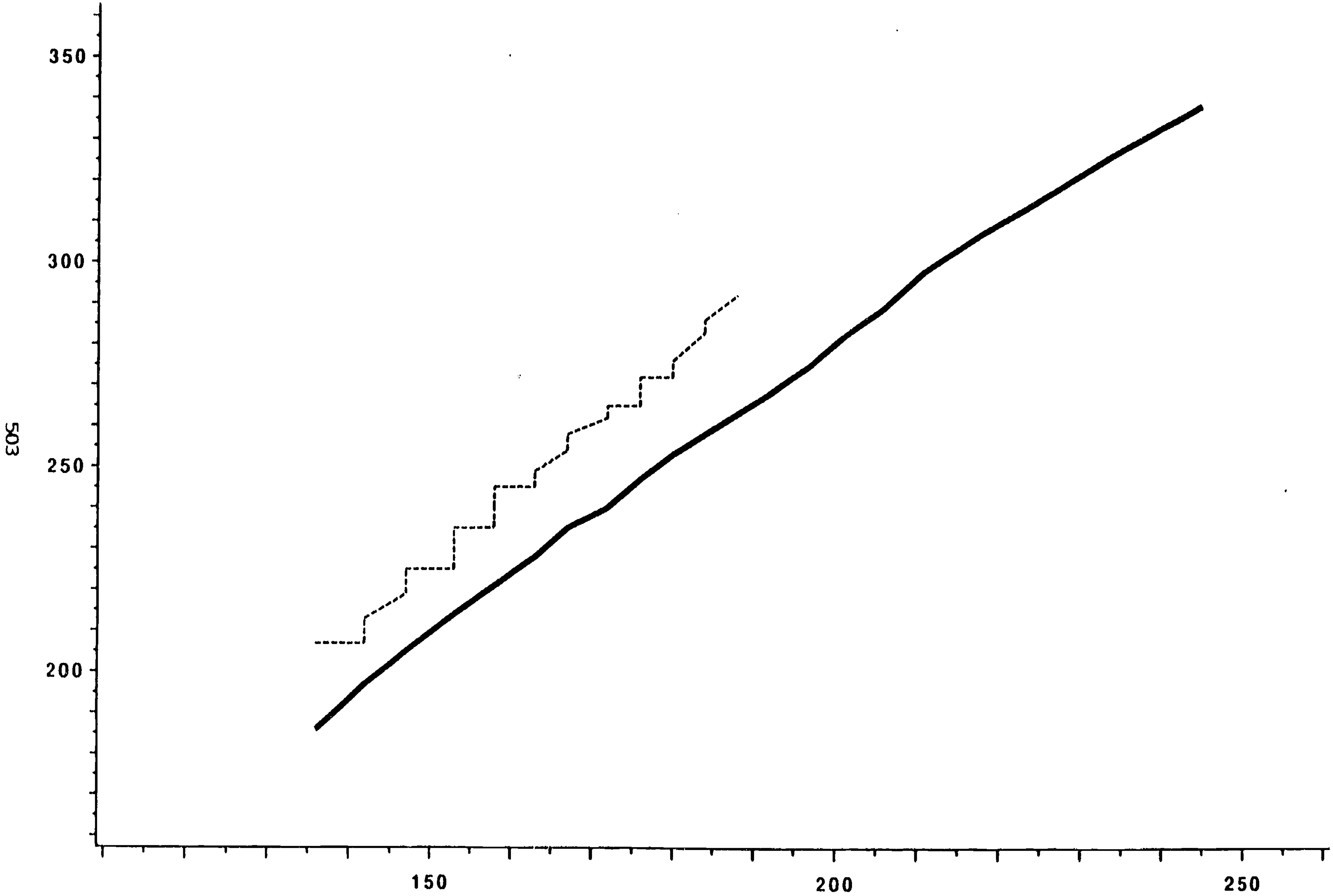
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

ERIC

Figure 135

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



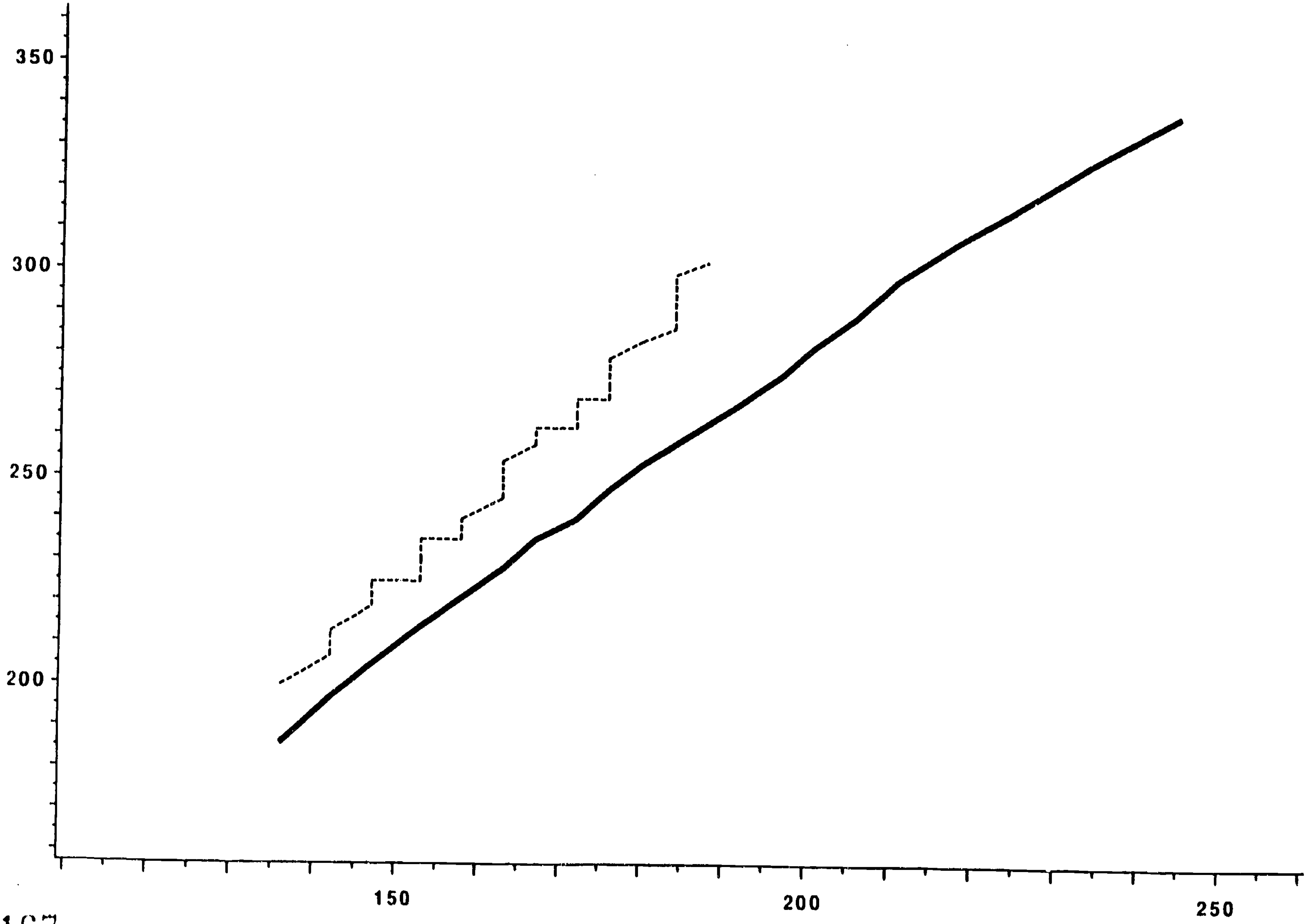
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 136

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H



CURVE ——— NORM -------- IS-H

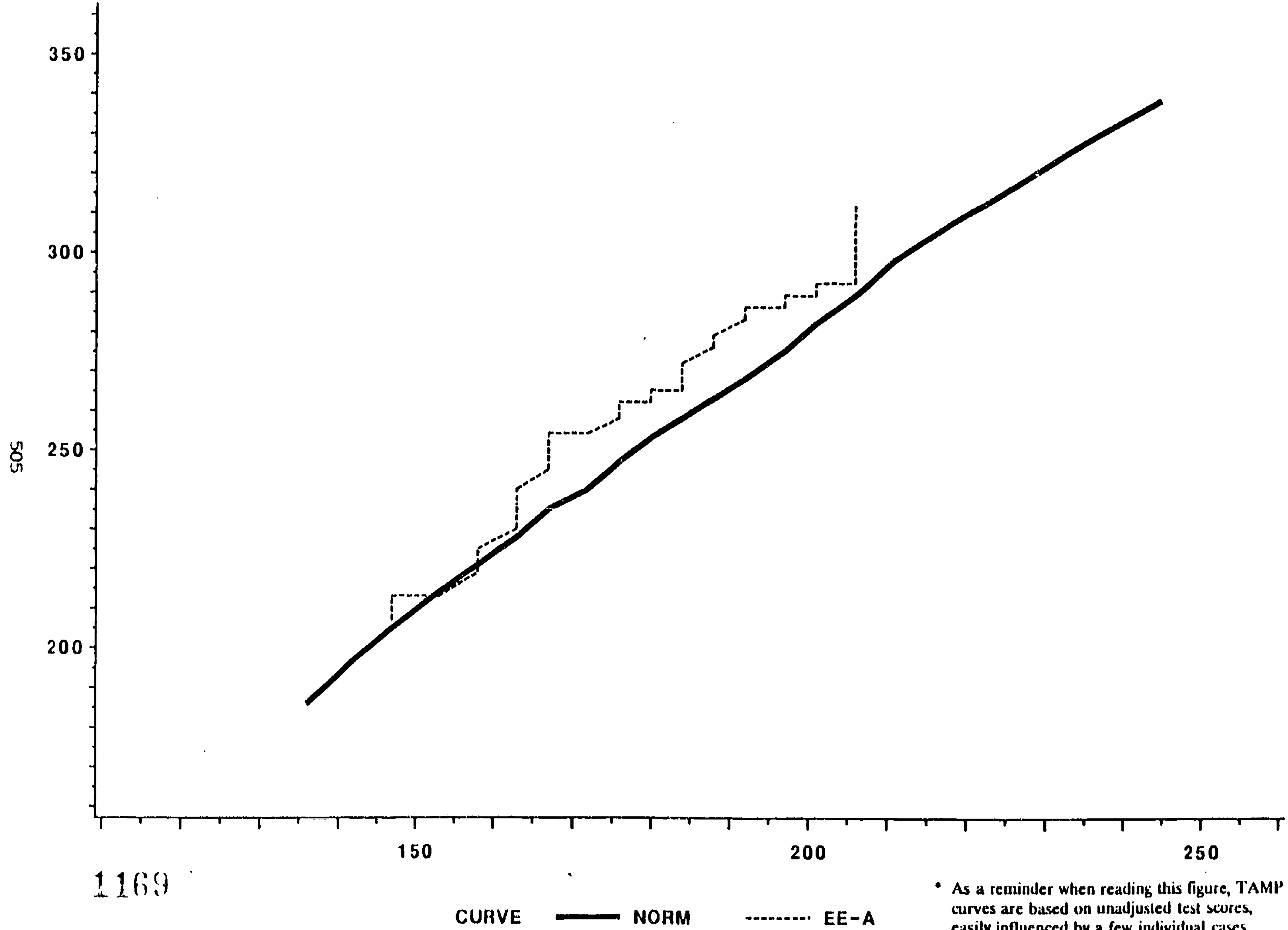
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

ERIC

Figure 137

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



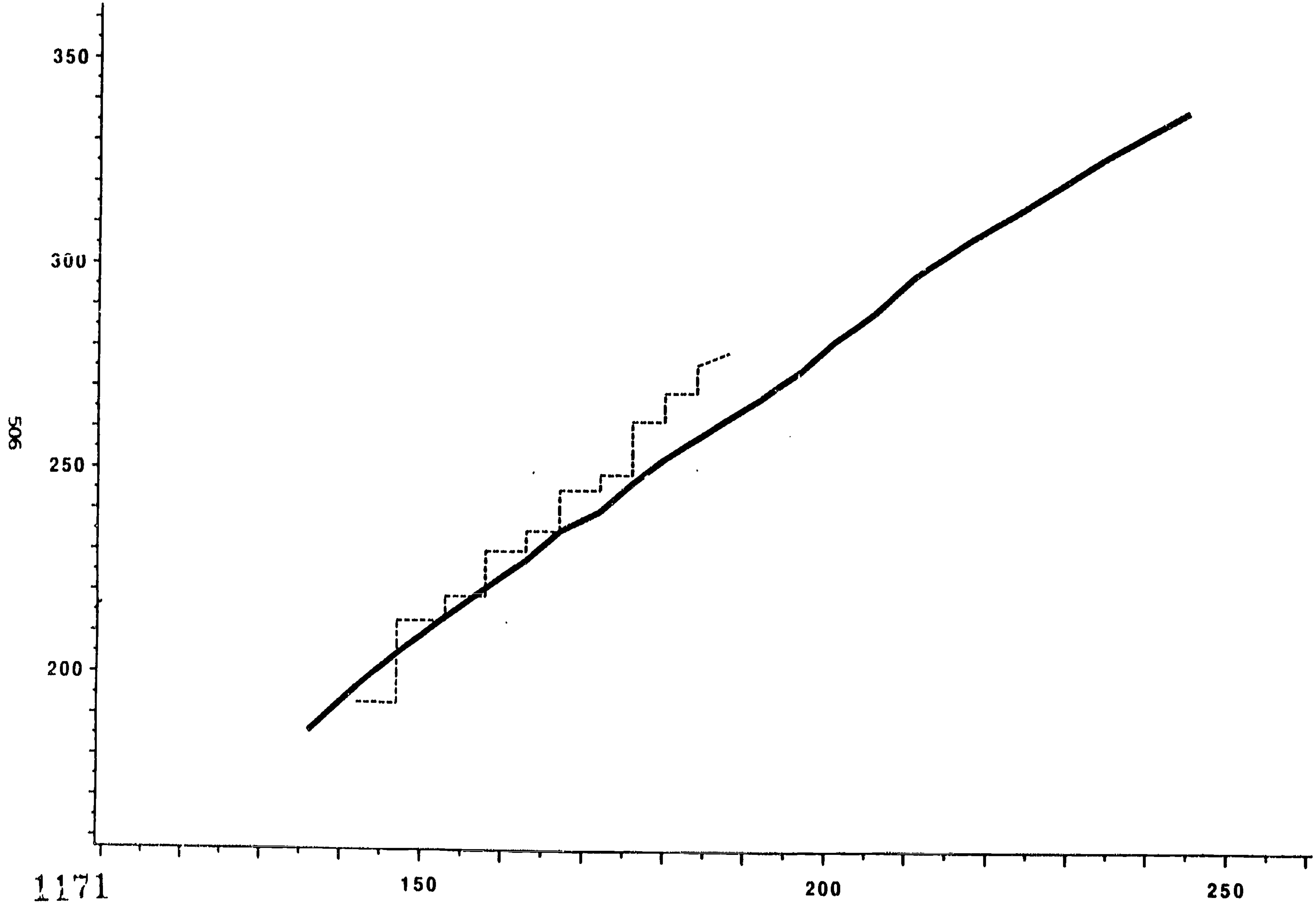
- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 138

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



CURVE ▬▬▬ NORM -------- EE-B

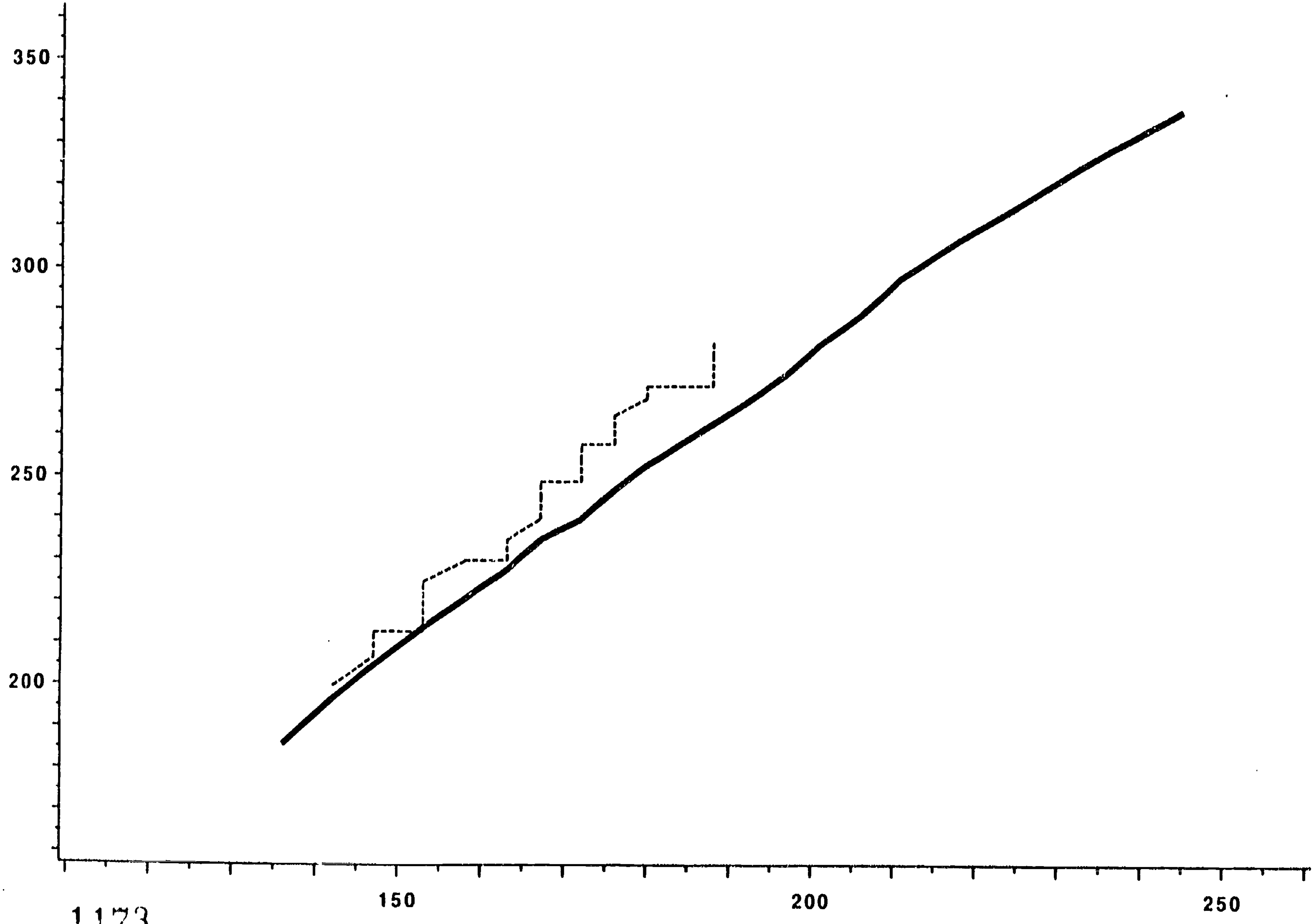
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

ERIC

Figure 139

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C

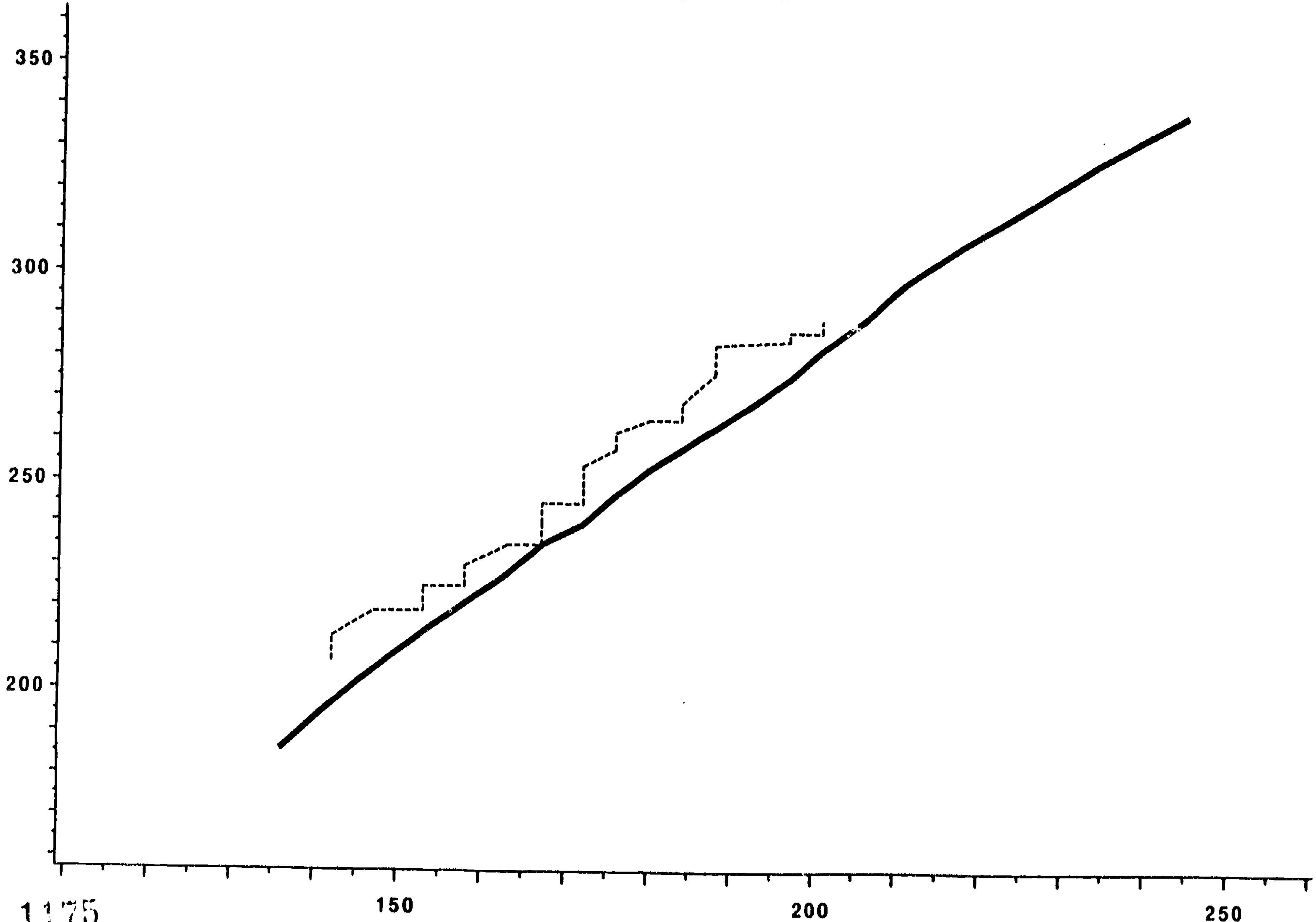


[a] As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 140

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



1175

CURVE ——— NORM -------- EE-H

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

1176

B.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Language**	**to**	**Language**

These analyses address two questions. First, how comparable were students in each of the three instructional programs with entry-level Spanish language skills relative to this norming population? Secondly, how comparable were these students in their English language skills at the end of first grade? The data suggest that immersion strategy, early-exit, and late-exit students began their kindergarten year with lower entry-level language skills than this norming population and that they all appear to have learned their English language skills at about the same rate as this norming population (see Figures 141, 142, and 143).

There is a modest suggestion that immersion strategy students with the highest entry-level primary language skills learned at a slightly faster rate relative to this norming population. This would be consistent with the underlying assumption of primary language programs that those students with stronger primary language skills are able to develop second language skills more quickly than students with lower primary language skills.

Late-exit students with lower entry-level language skills appear to have learned English language skills at about the same rate as comparable students in this norming population. At first glance, this seems somewhat surprising given the limited amount of instruction provided in English. However, given the very basic skills that these low achievement levels represent, increases in English skills probably reflect the exposure to English that these students receive in their out-of-classroom environment, e.g., television, playground, etc. The lower growth rate in English language skills of late-exit students with the highest entry-level Spanish language skills relative to this norming population is consistent with the predictions made from the late-exit instructional model, i.e., given less instruction in English, slower growth in English skills would result.

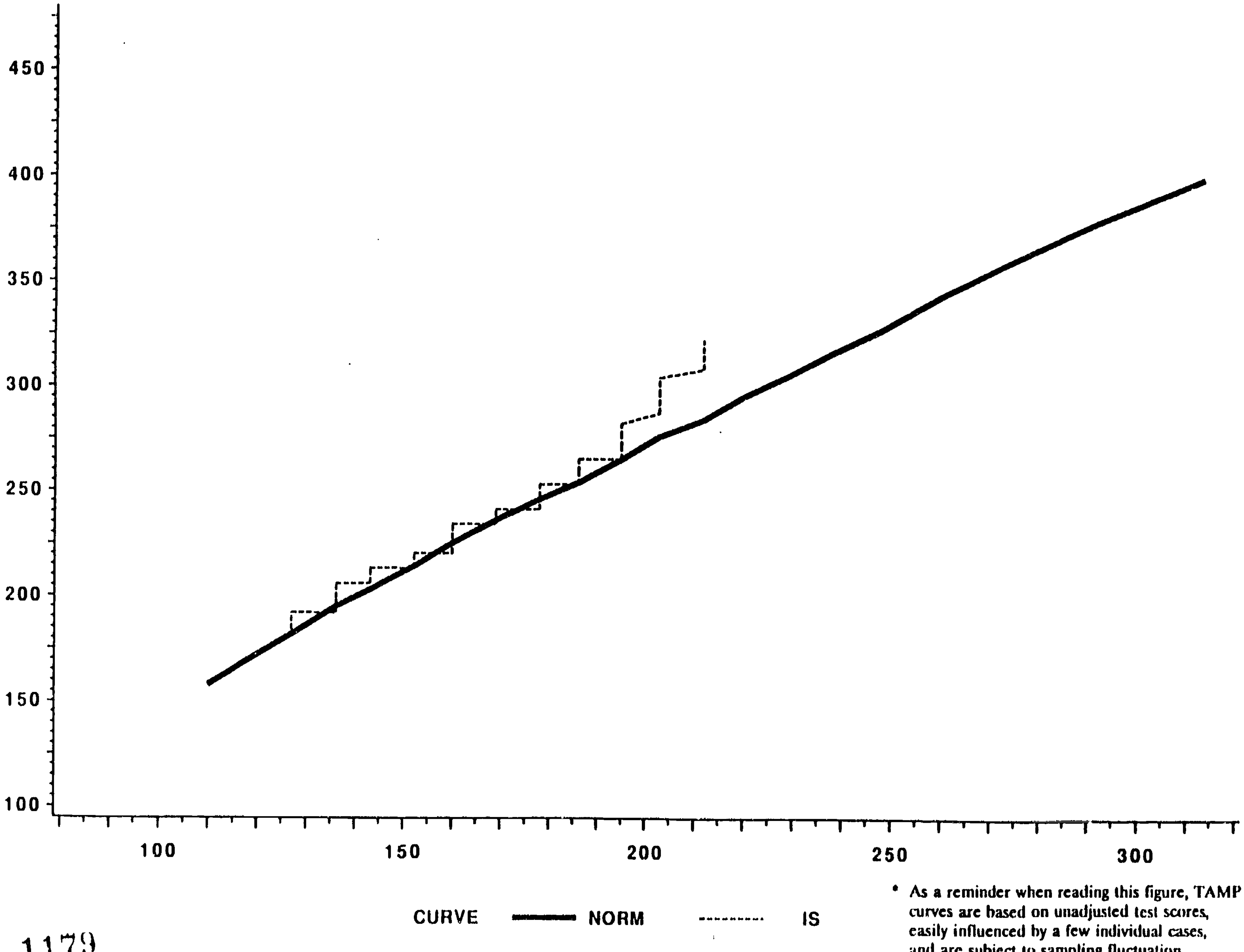
As before, when the TAMP curves for each program site are examined, little variation was noted for either the immersion strategy (see Figures 144 to 147) or the early-exit program (see Figures 148 to 151). The immersion strategy students with the highest entry-level scores were mostly in site IS-H, so that this district accounts for the higher growth rate among these students.

In sum, students in all three programs appear to enter with comparable primary language skills, and have generally comparable English language skills at the end of first grade. Noteworthy is that there is some evidence in support of the notion that students with higher entry-level skills in their primary language acquire second language skills at a faster rate than students with lower first language skills.

Figure 141

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

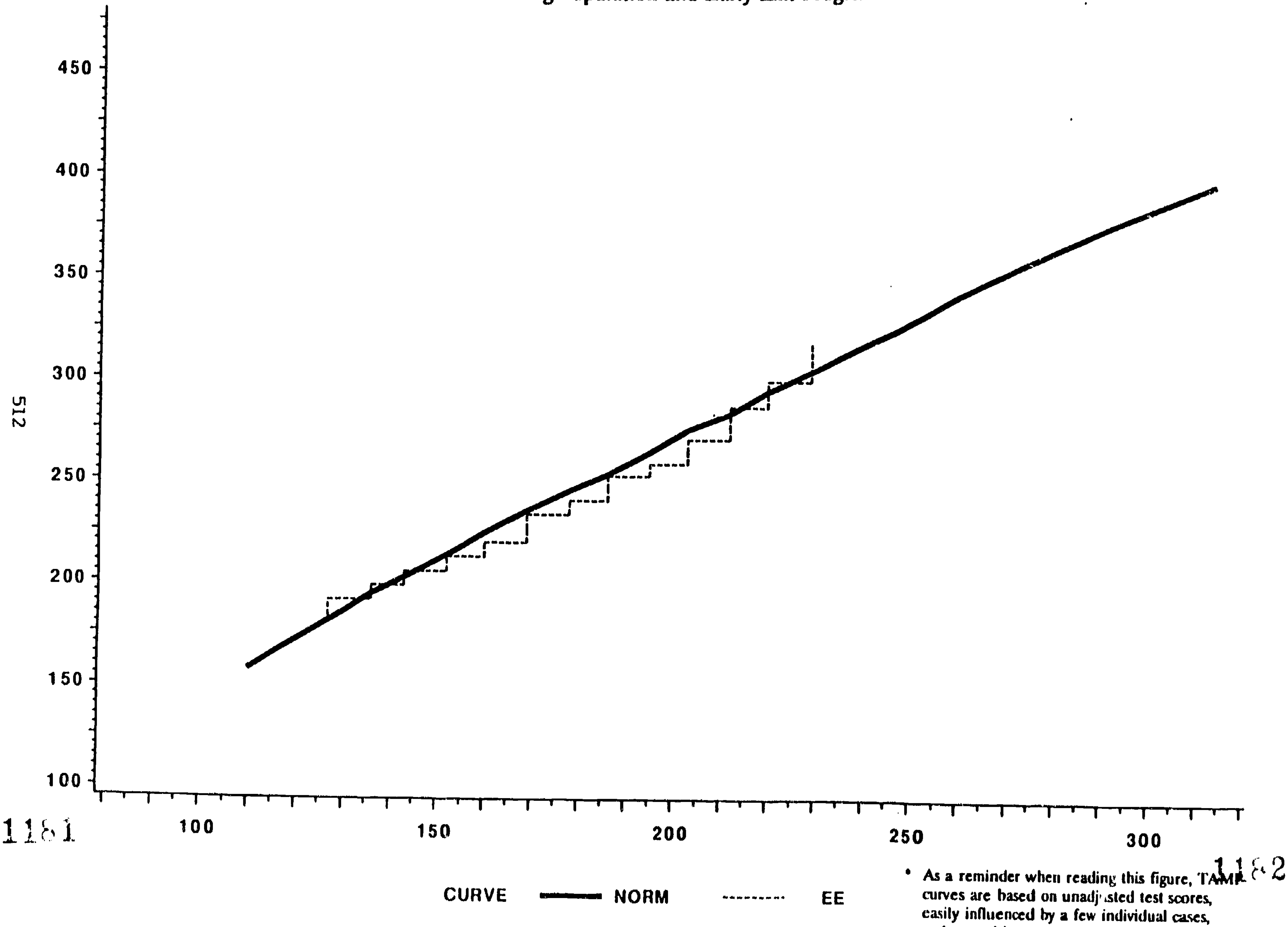


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 142

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

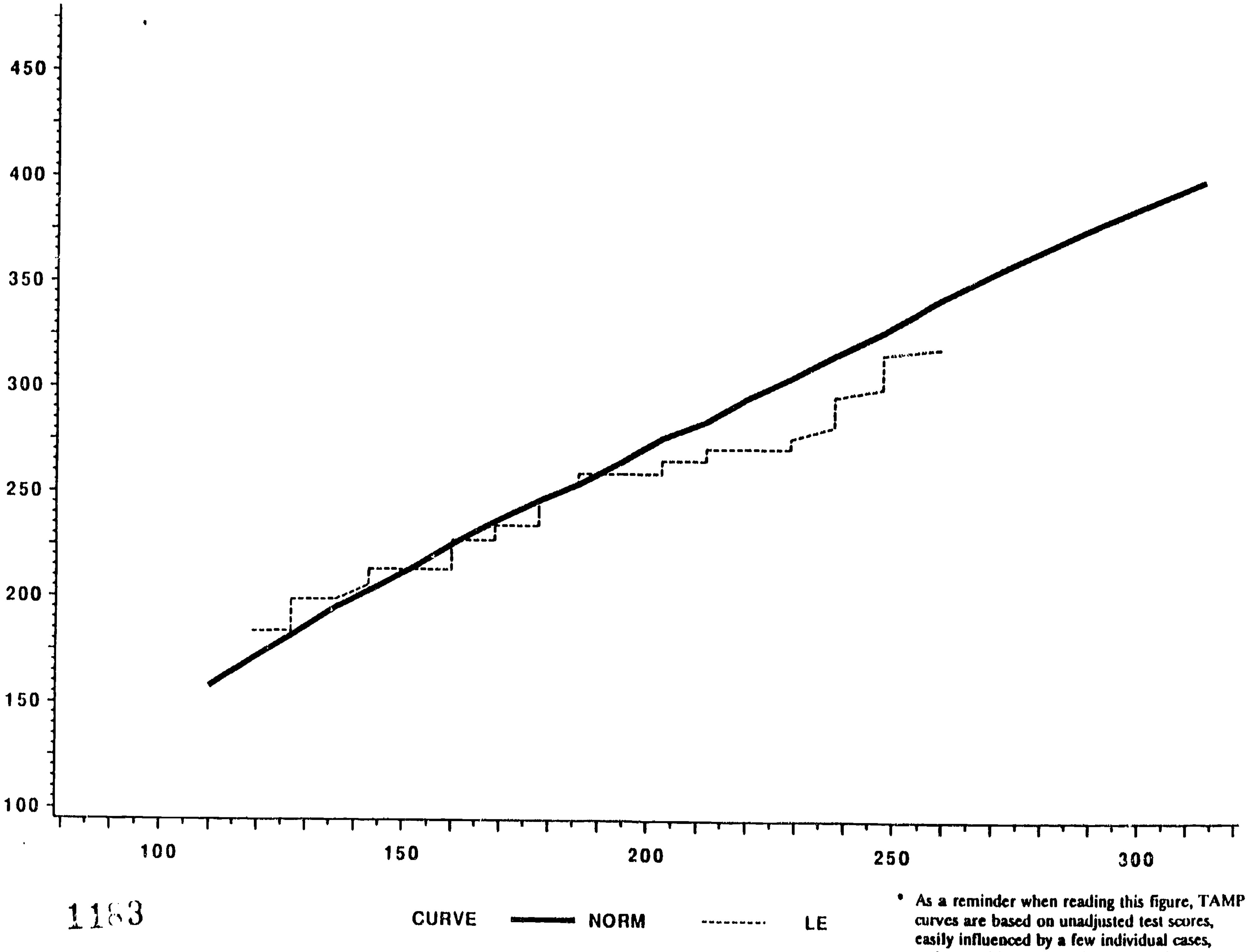


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 143

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



513

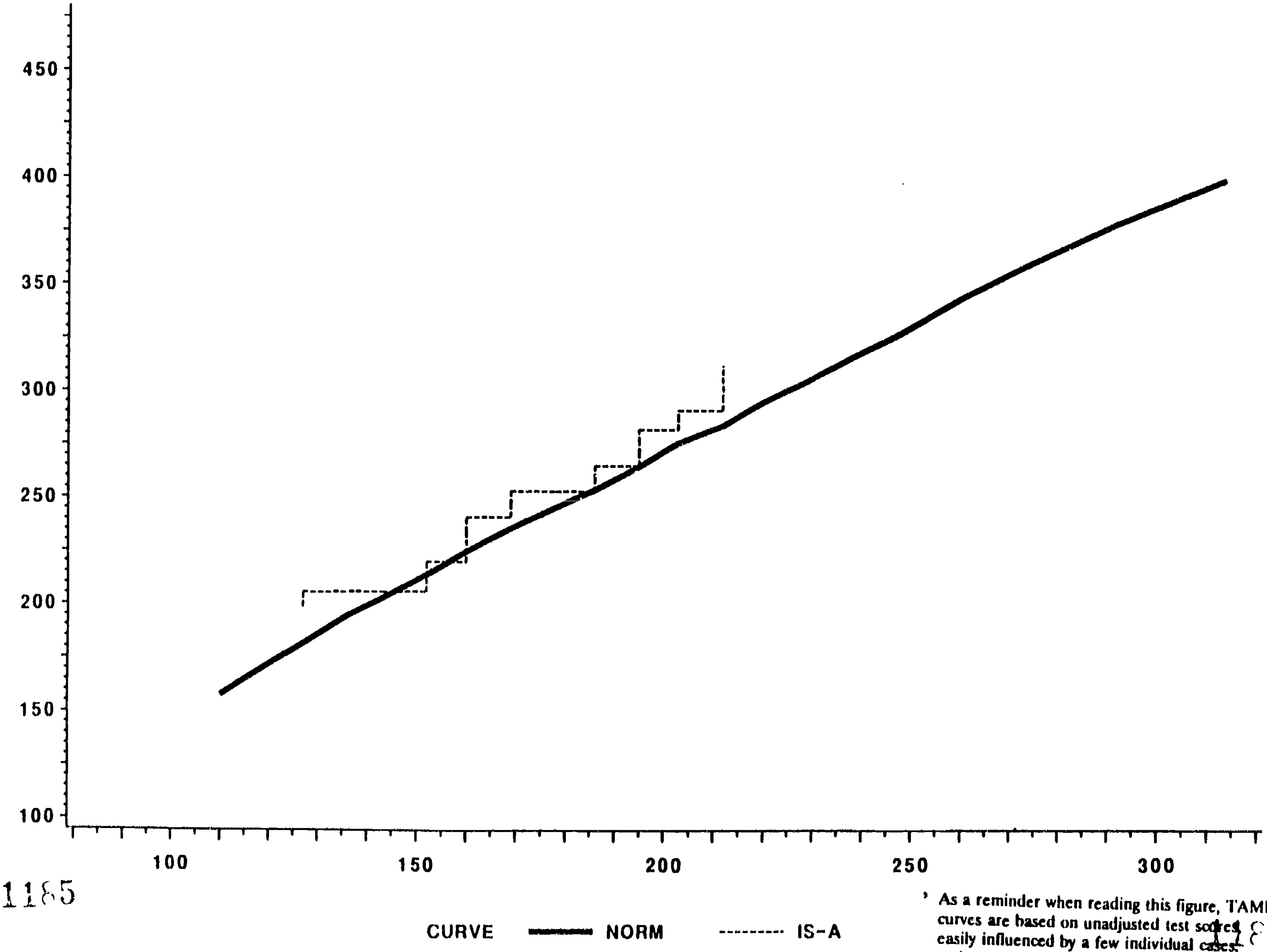
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 144

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



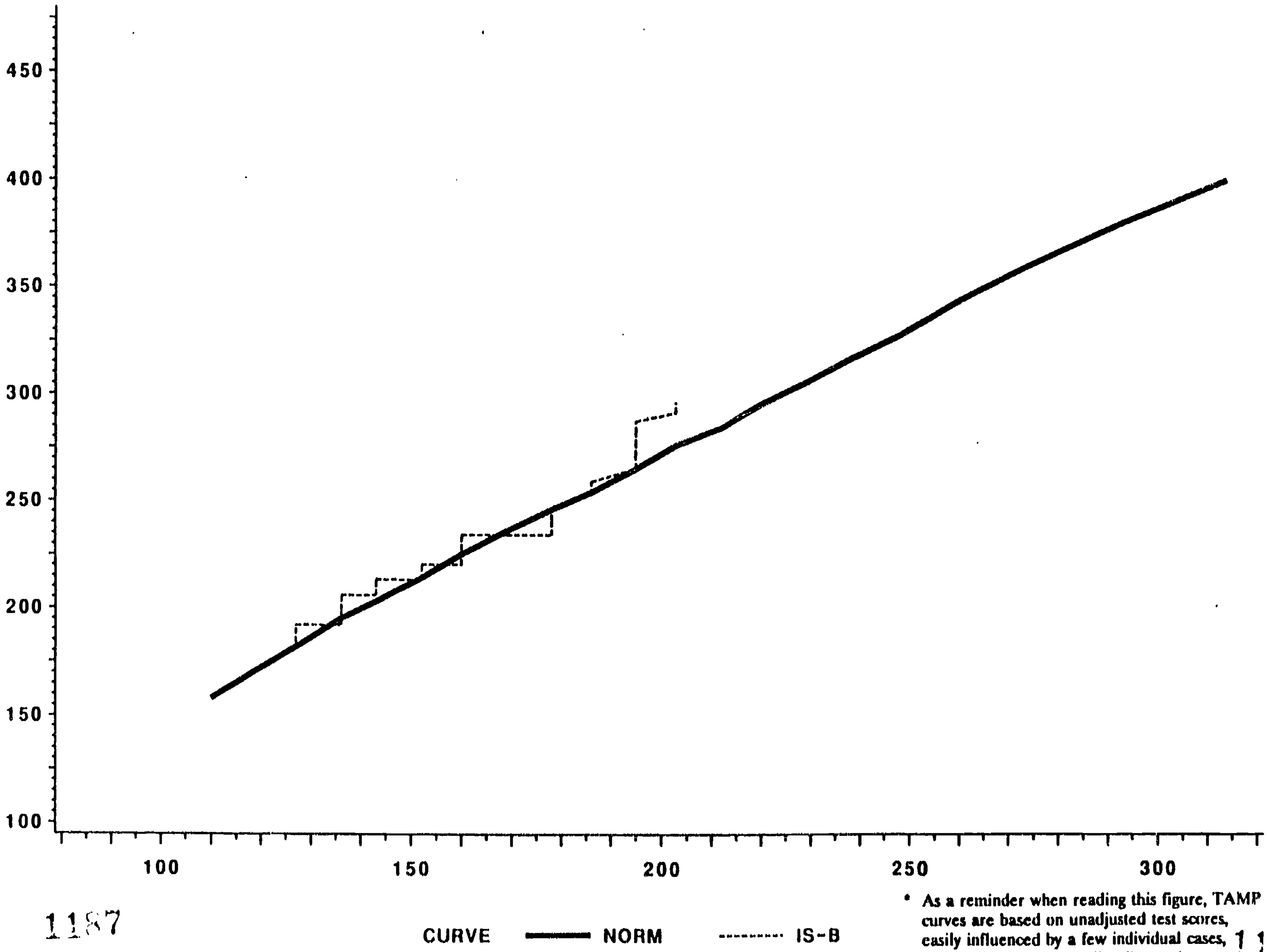
' As a reminder when reading this figure, TAMP curves are based on unadjusted test scores easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 145

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



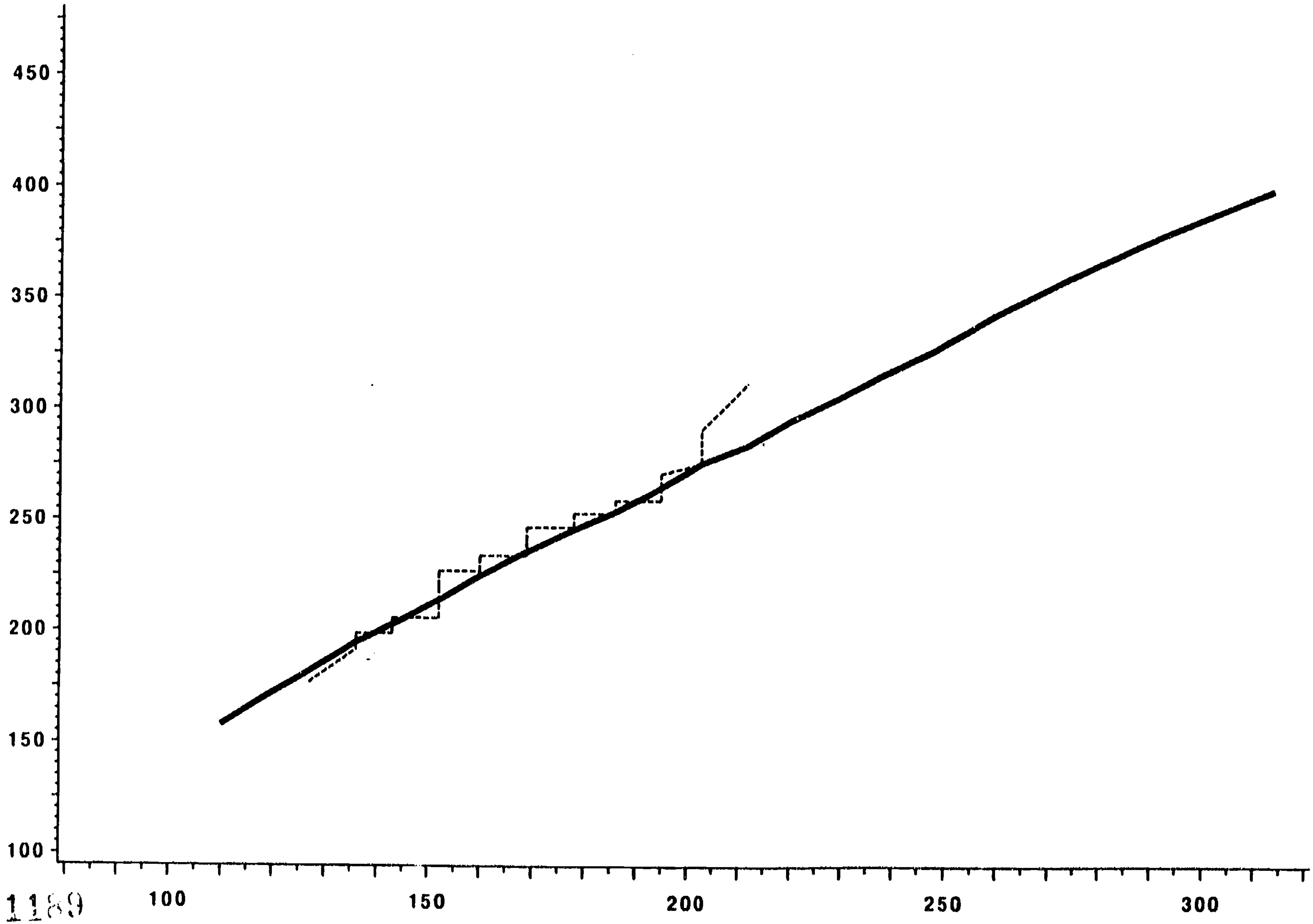
515

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 146

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



516

1189

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

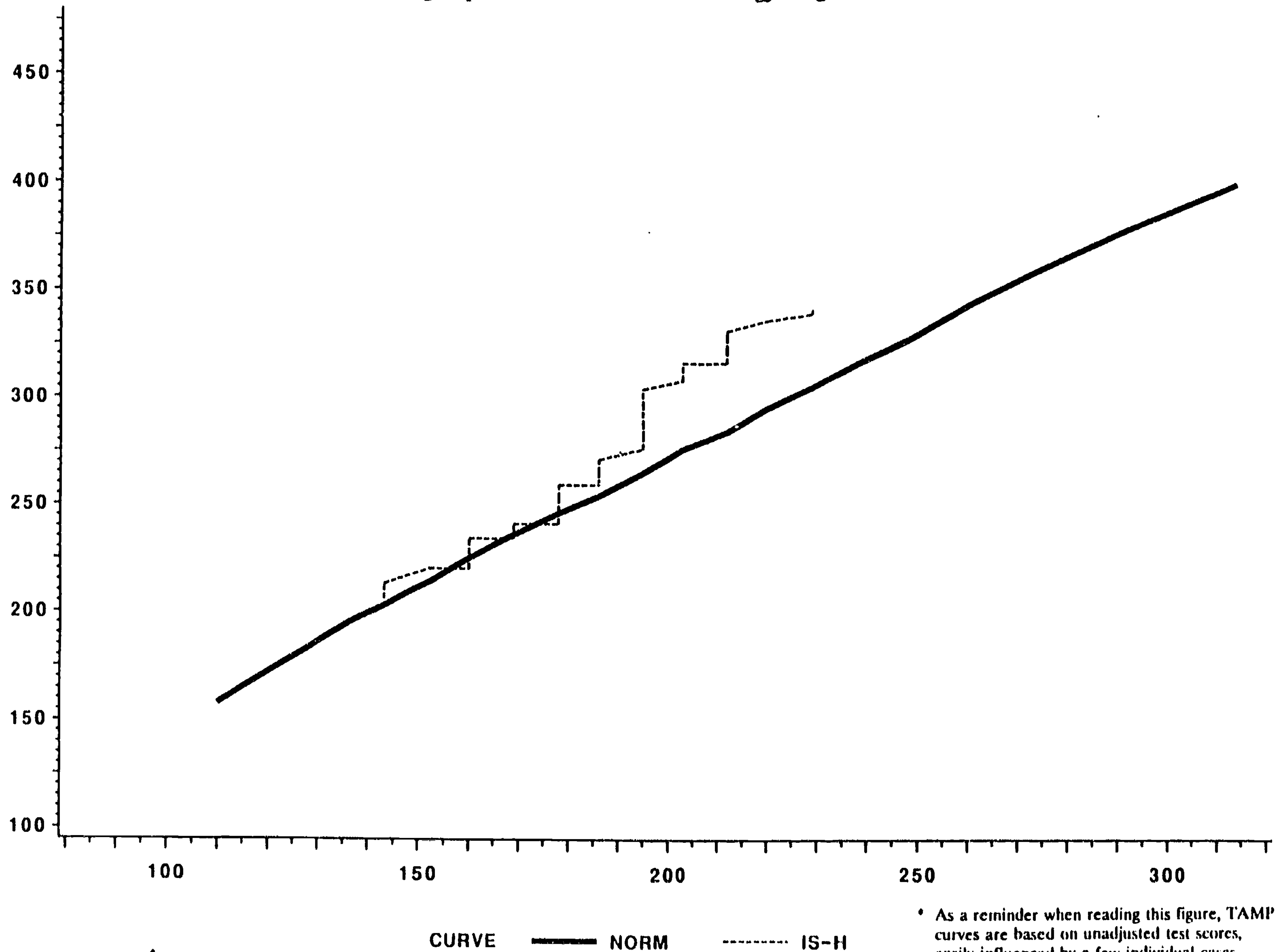
1190

ERIC

Figure 147

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H

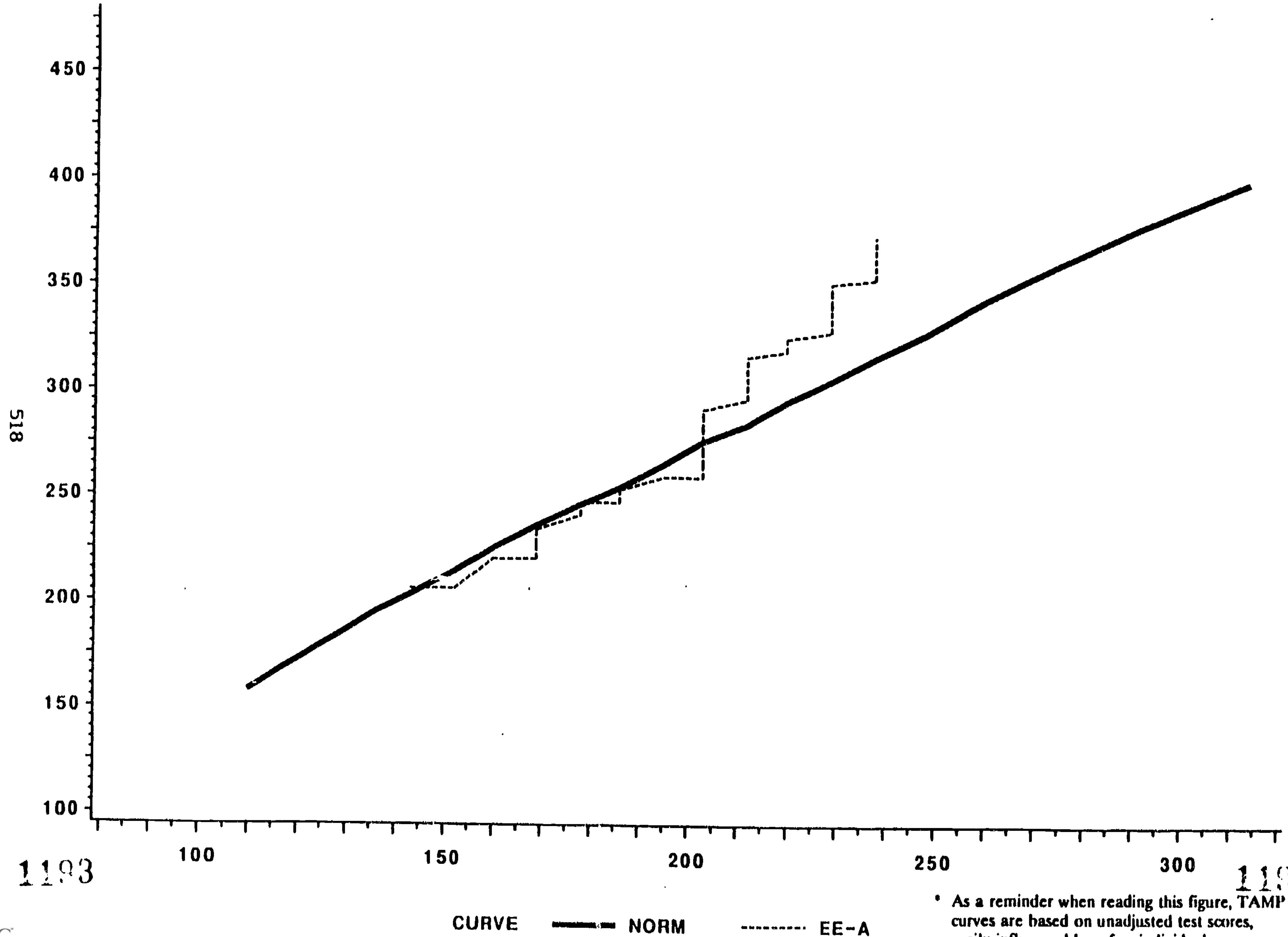


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 148

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



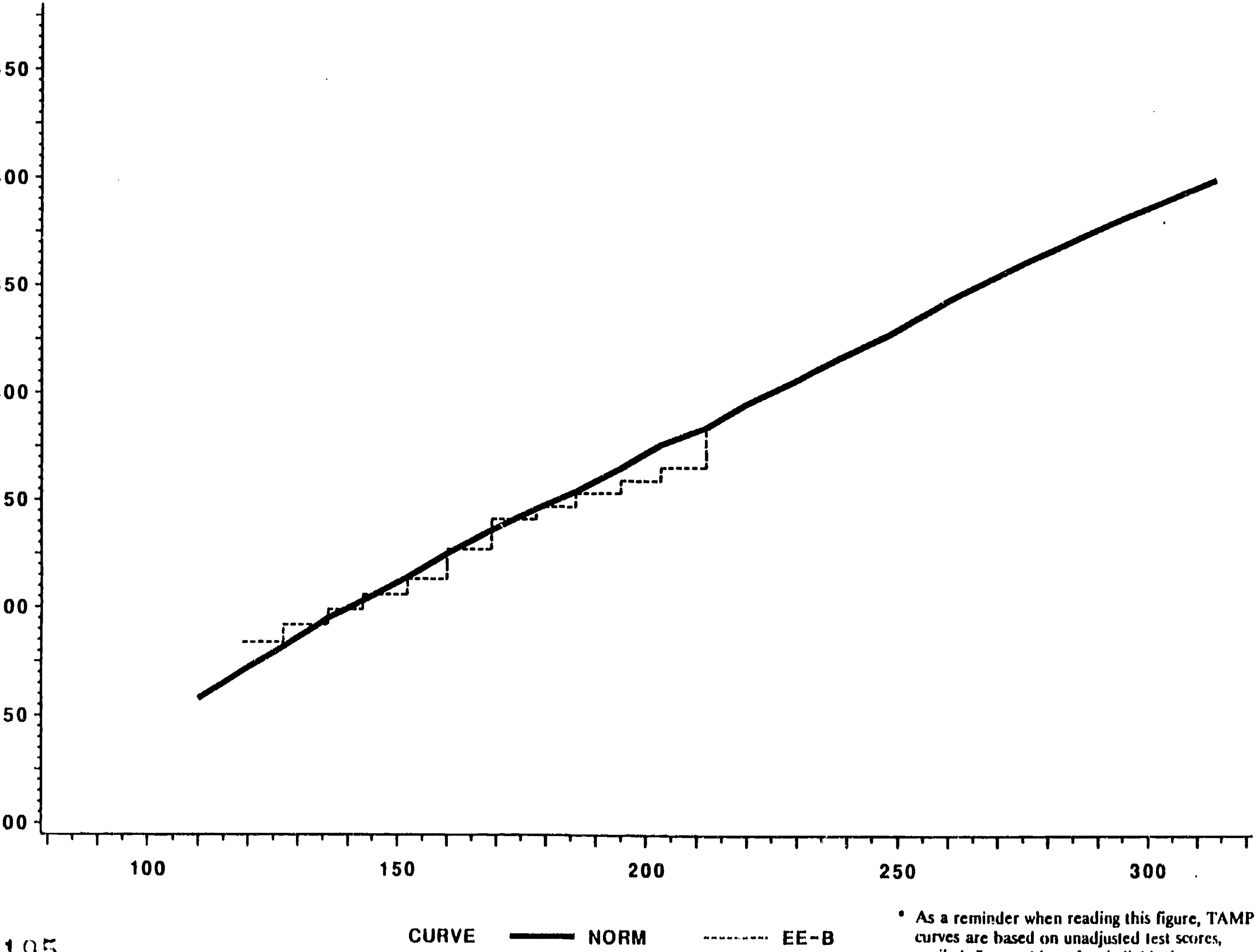
1193 1194

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 149

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B



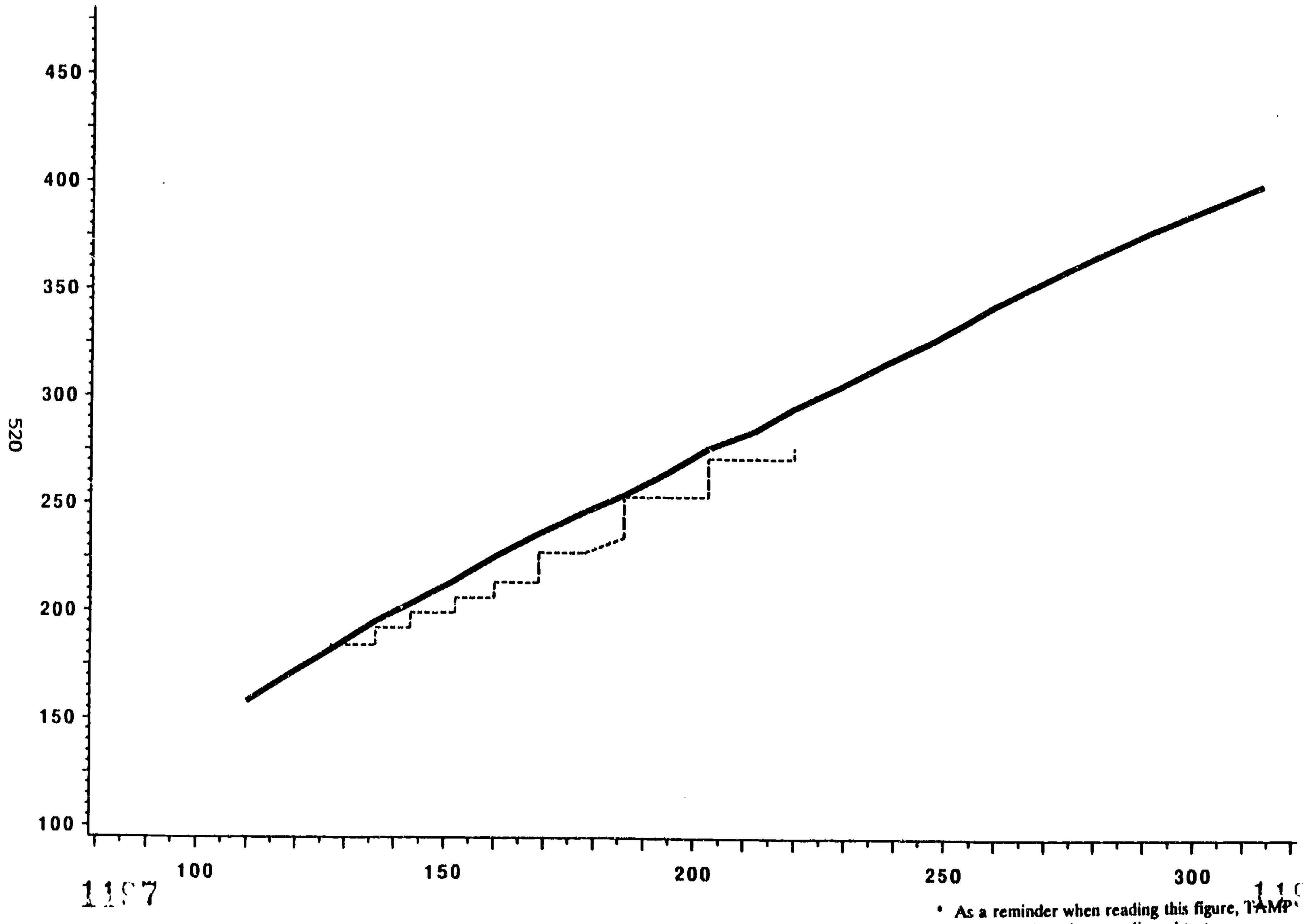
519

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 150

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



1197 1198

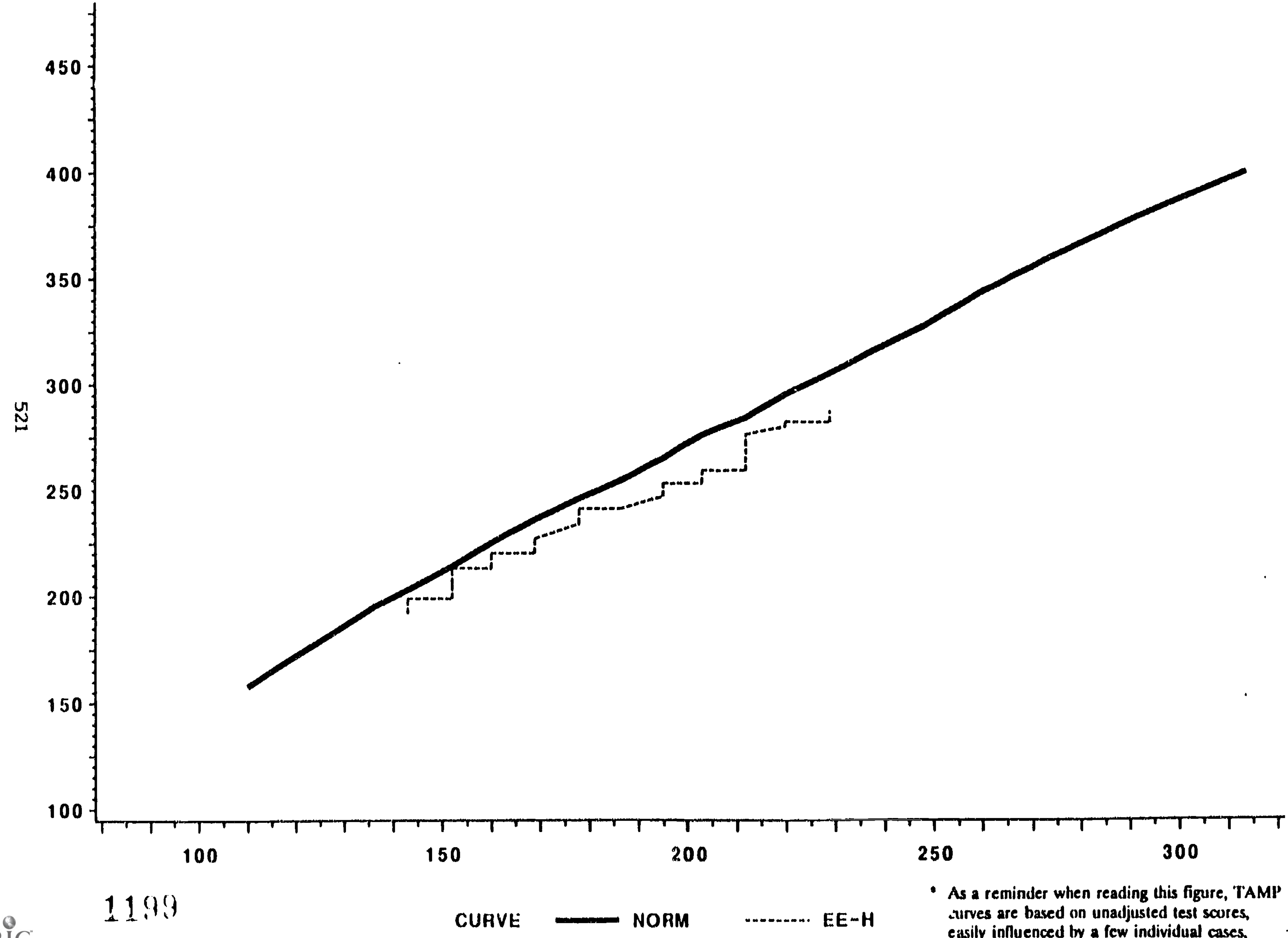
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

ERIC

Figure 151

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

C.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Language**	**to**	**Reading**

It is hypothesized that students who have higher primary language skills are able to acquire second language skills more easily. These analyses attempt to determine the extent to which immersion strategy, early-exit, and late-exit students have comparable entry-level skills relative to this norming population, and to determine if the growth rate in English reading for each of these three groups is comparable to that of this norming population.

Once again both immersion strategy and early-exit students appear to be growing in English reading skills at the same rate as this norming population and as each other (see Figures 152 and 153). As predicted by the late-exit model, the achievement in English reading for late-exit students does not seem to be comparable to that of this norming population (see Figure 154) or to either the immersion strategy or early-exit students. This is considered reasonable as late-exit students generally did not receive instruction in English reading in kindergarten or first grade. Late-exit students who entered kindergarten with low Spanish language skills appear to have increased their English reading skills relative to that of this norming population. As before, this increase in English reading skills by students with lower entry-level skills presumably reflects the most basic English reading skills which may have resulted from exposure to English in their environment.

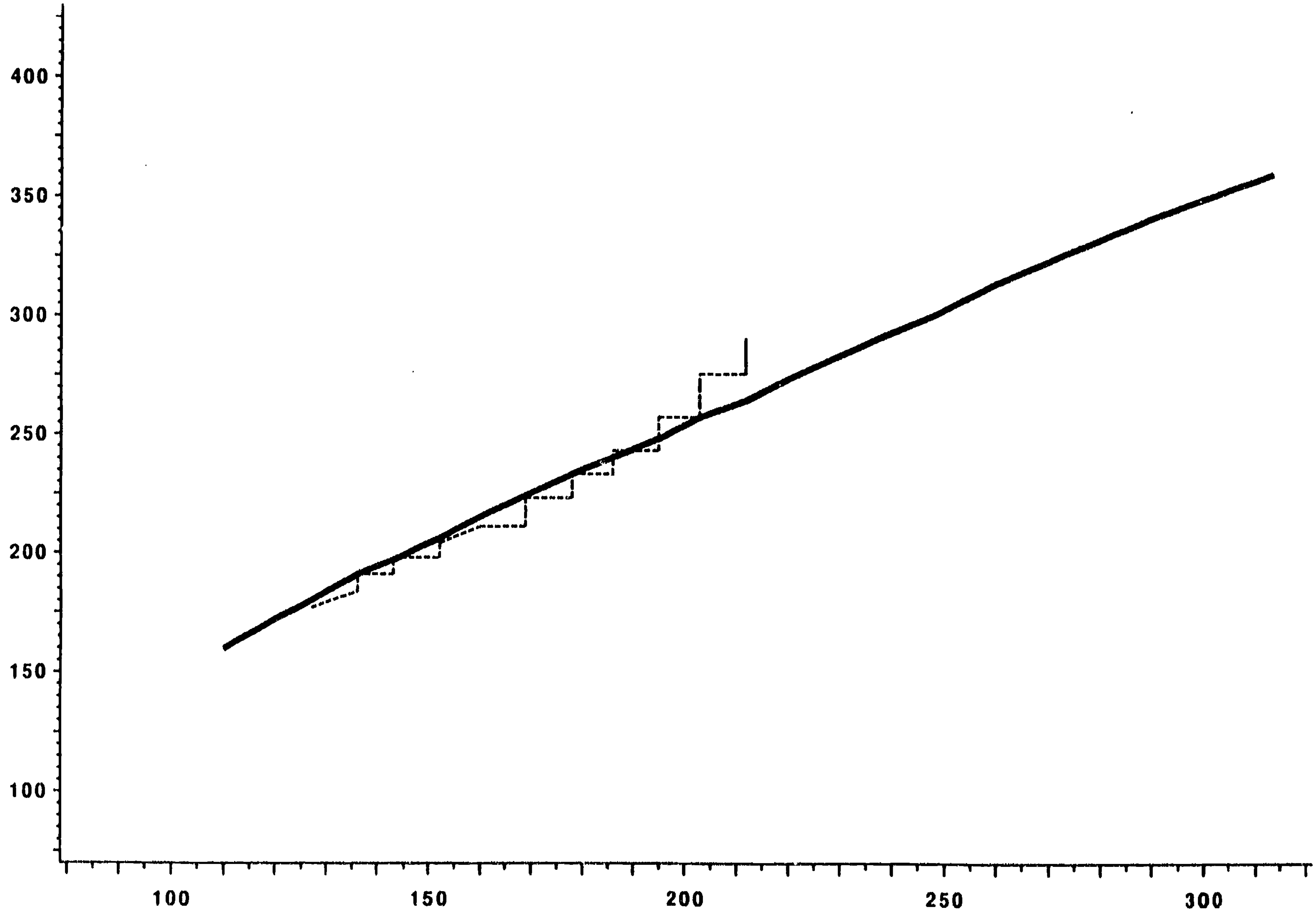
With the exception of site IS-H, little variation was noted among the individual immersion strategy (see Figures 155 to 158) or early-exit sites (see Figures 159 to 162). A similar growth rate had been predicted for immersion strategy and early-exit students. As before, the growth rate for site IS-H is most responsible for the composite immersion strategy site TAMP curve approximating the growth rate of this norming population.

ERIC

Figure 152

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program

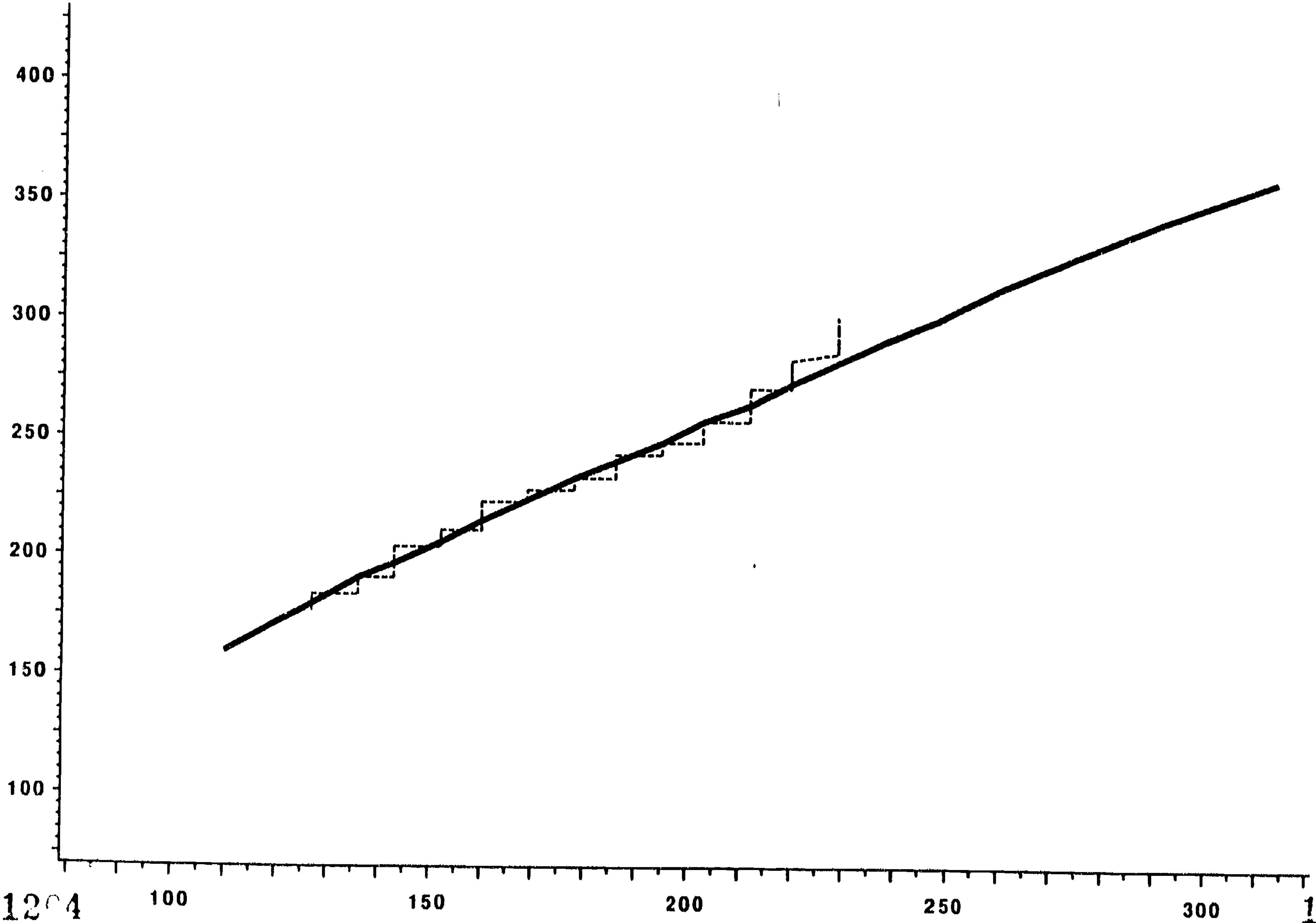


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 153

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program



CURVE ▬▬▬ NORM ------- EE

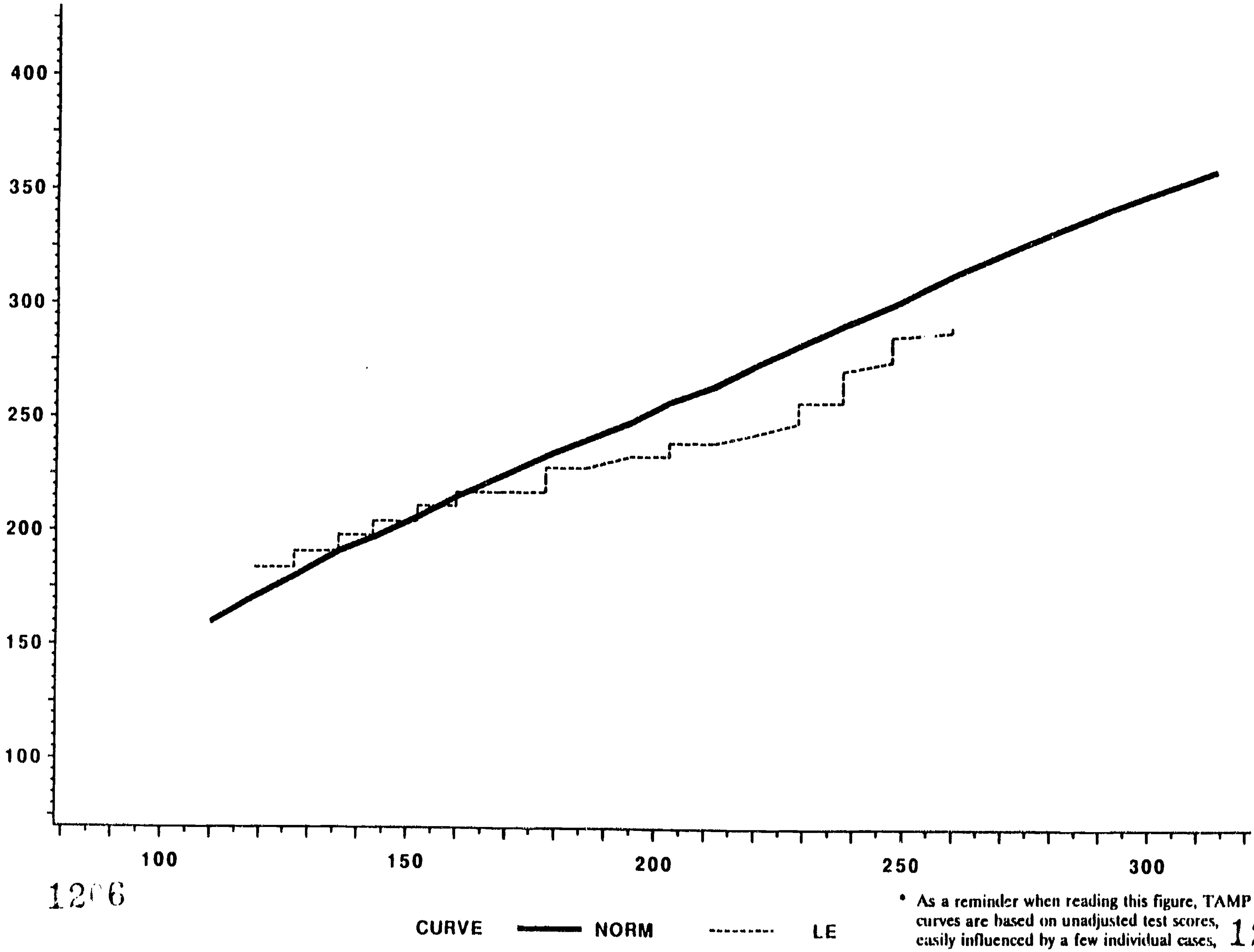
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

Figure 154

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



525

1206

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

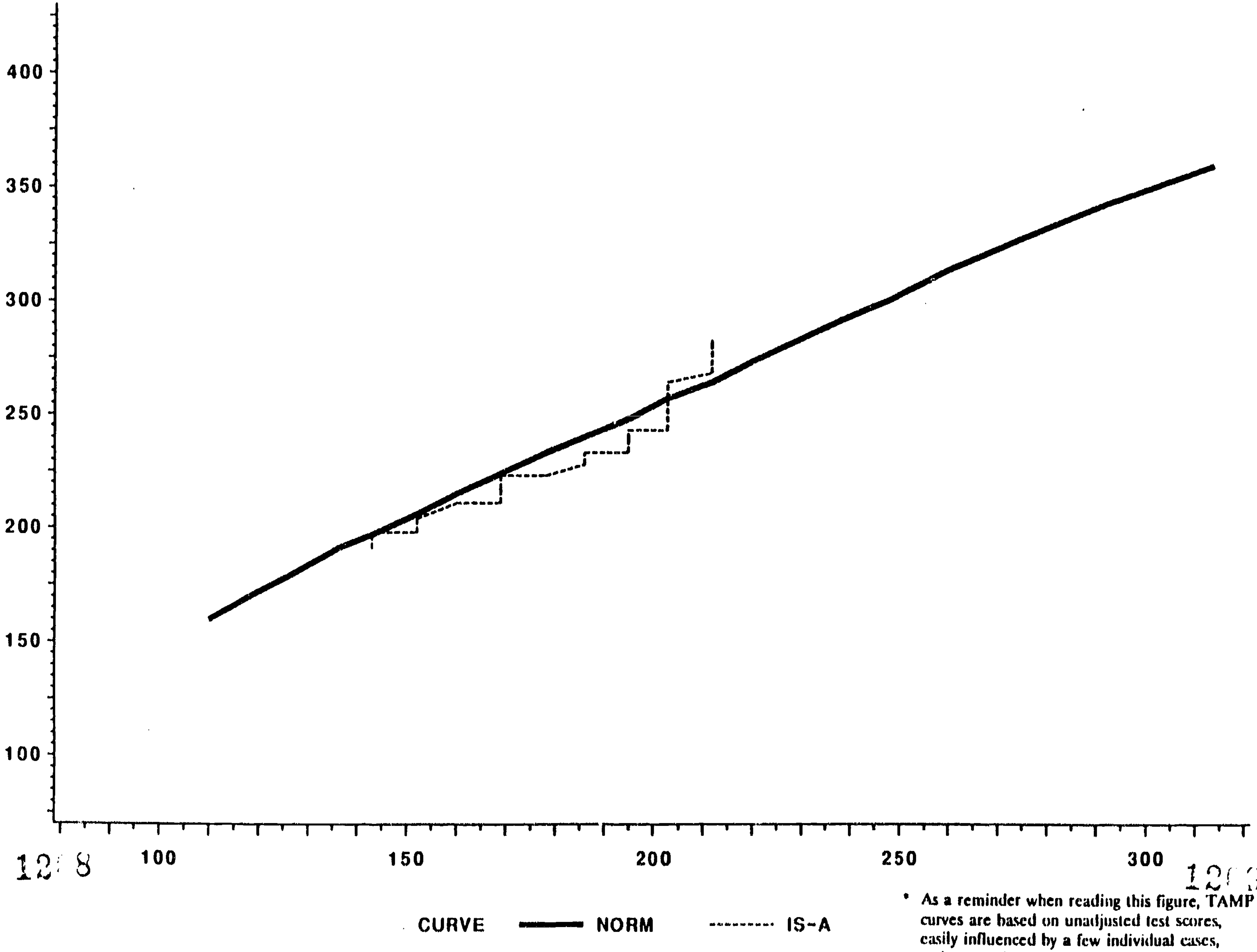
1207

ERIC

Figure 155

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District A



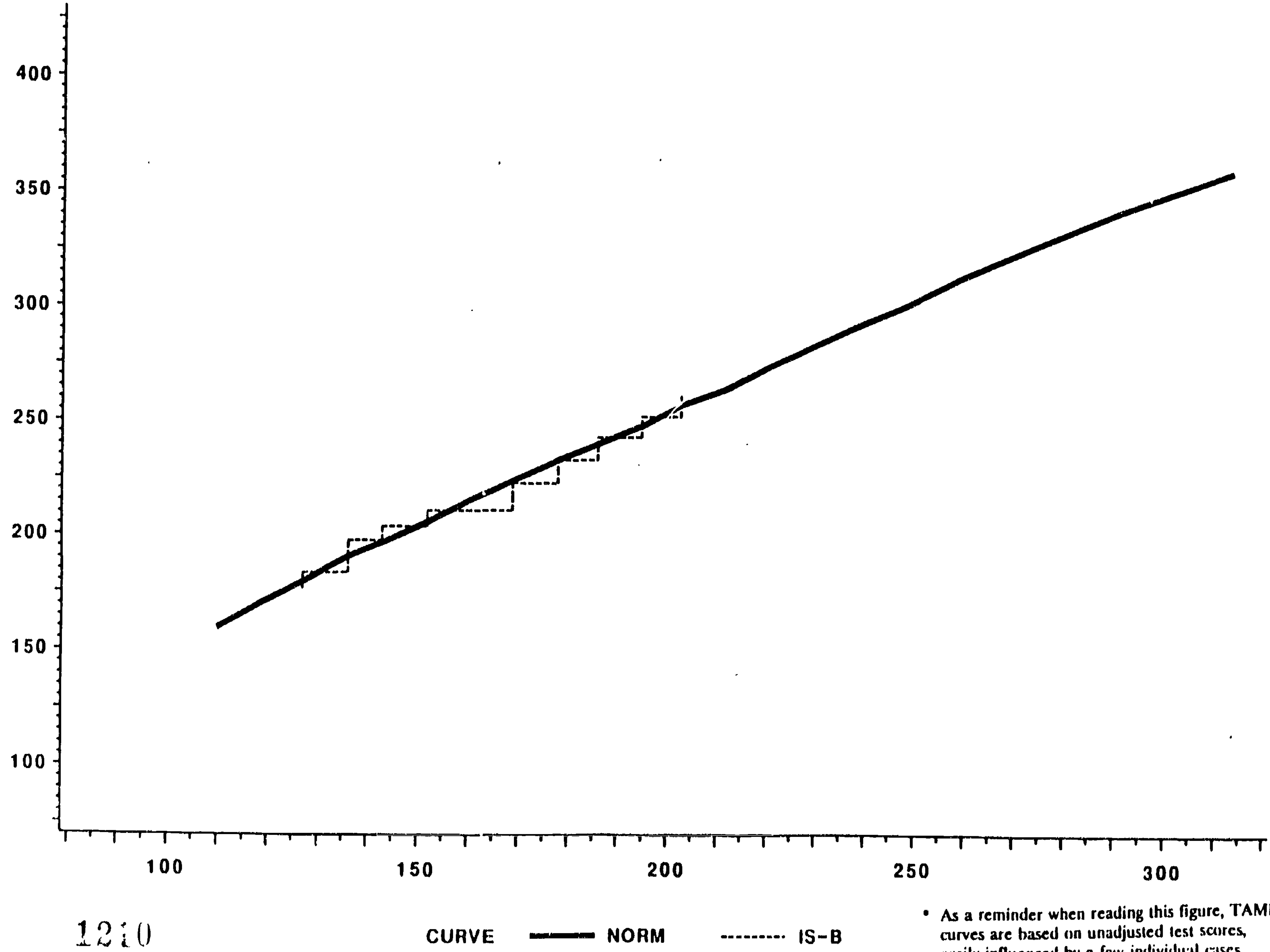
526

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 156

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District B



527

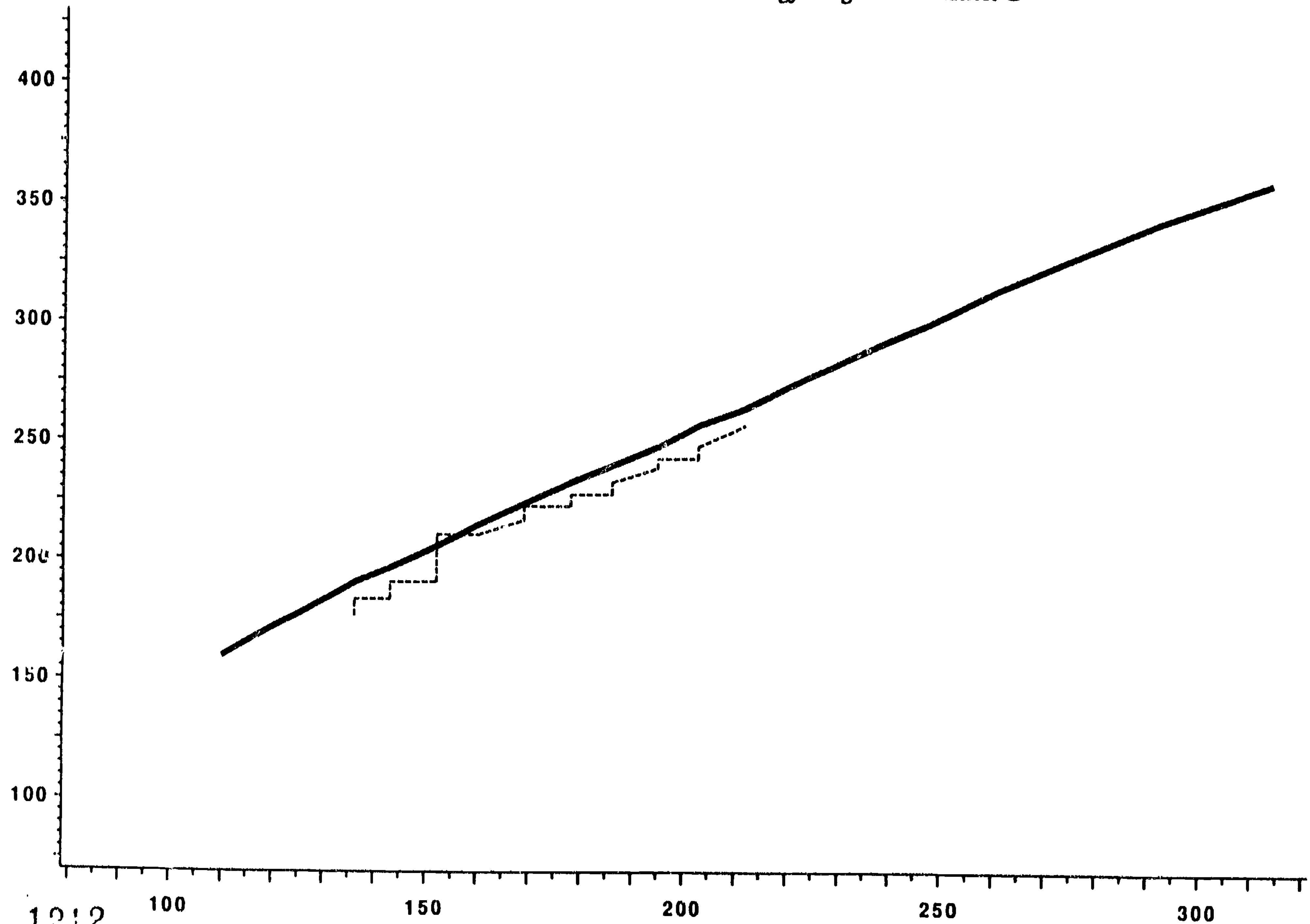
1210

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 157

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District C



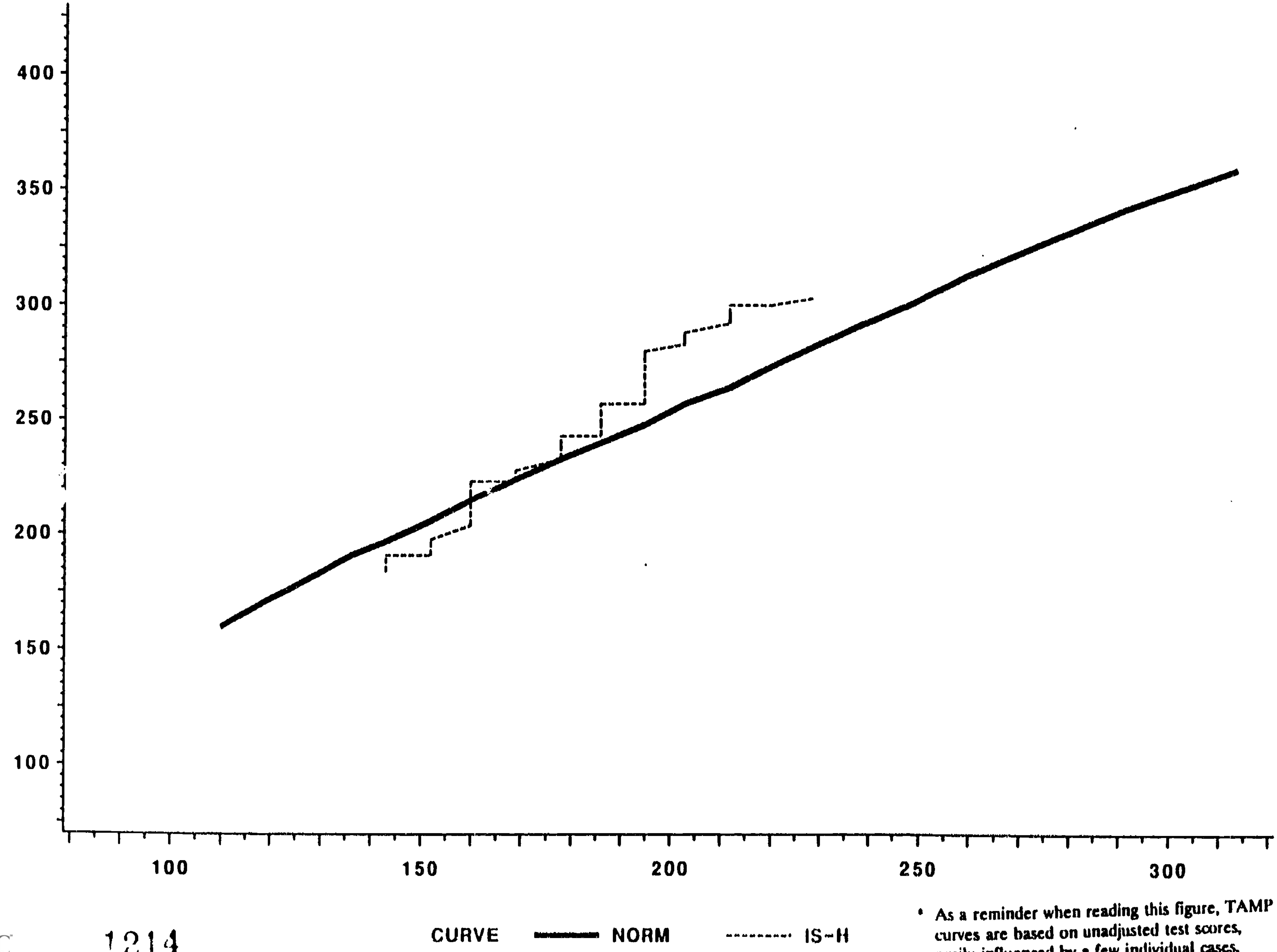
CURVE —— NORM ------- IS-C

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 158

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program in District H



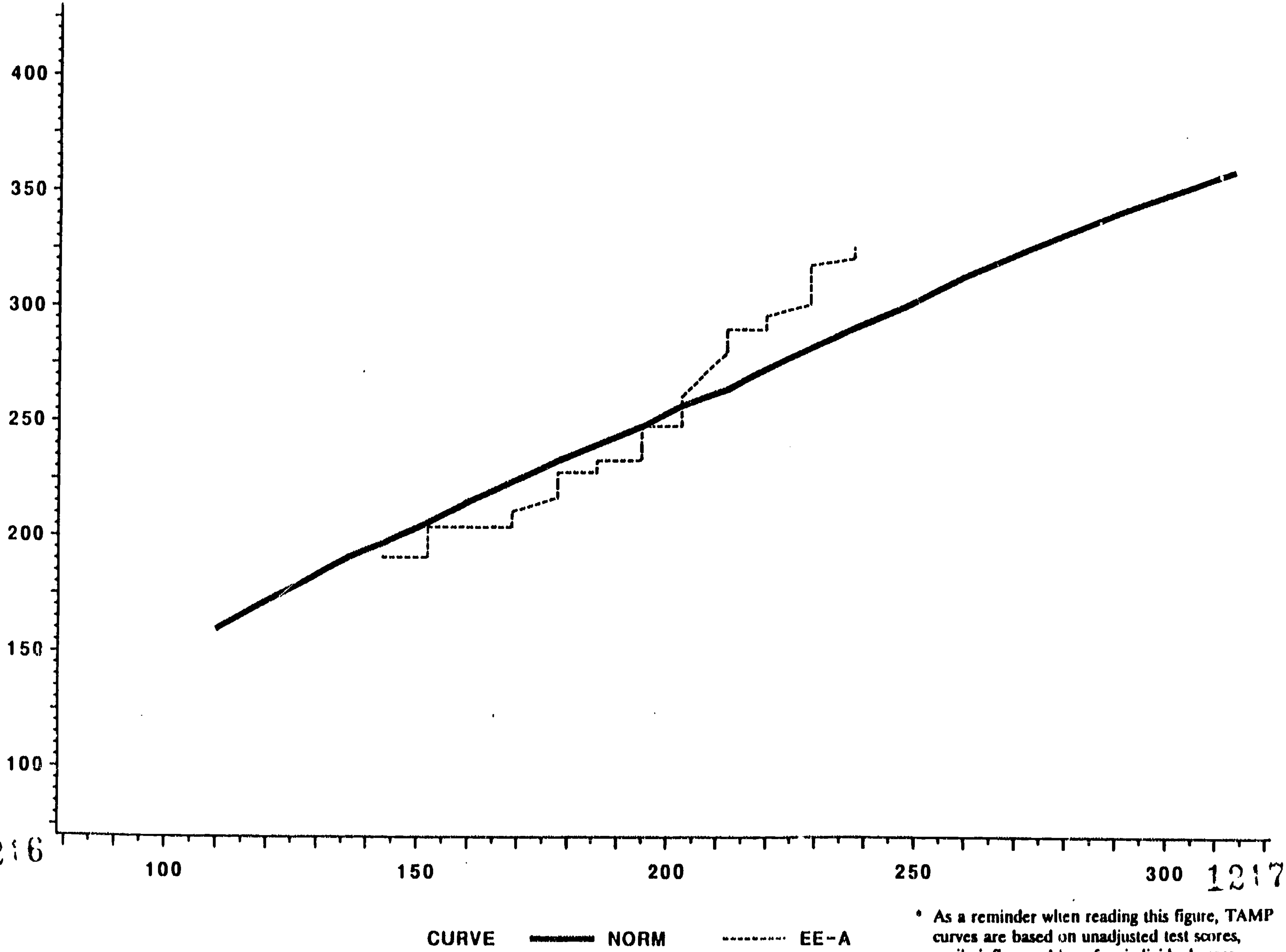
• As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

529

Figure 159

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District A



530

1216 1217

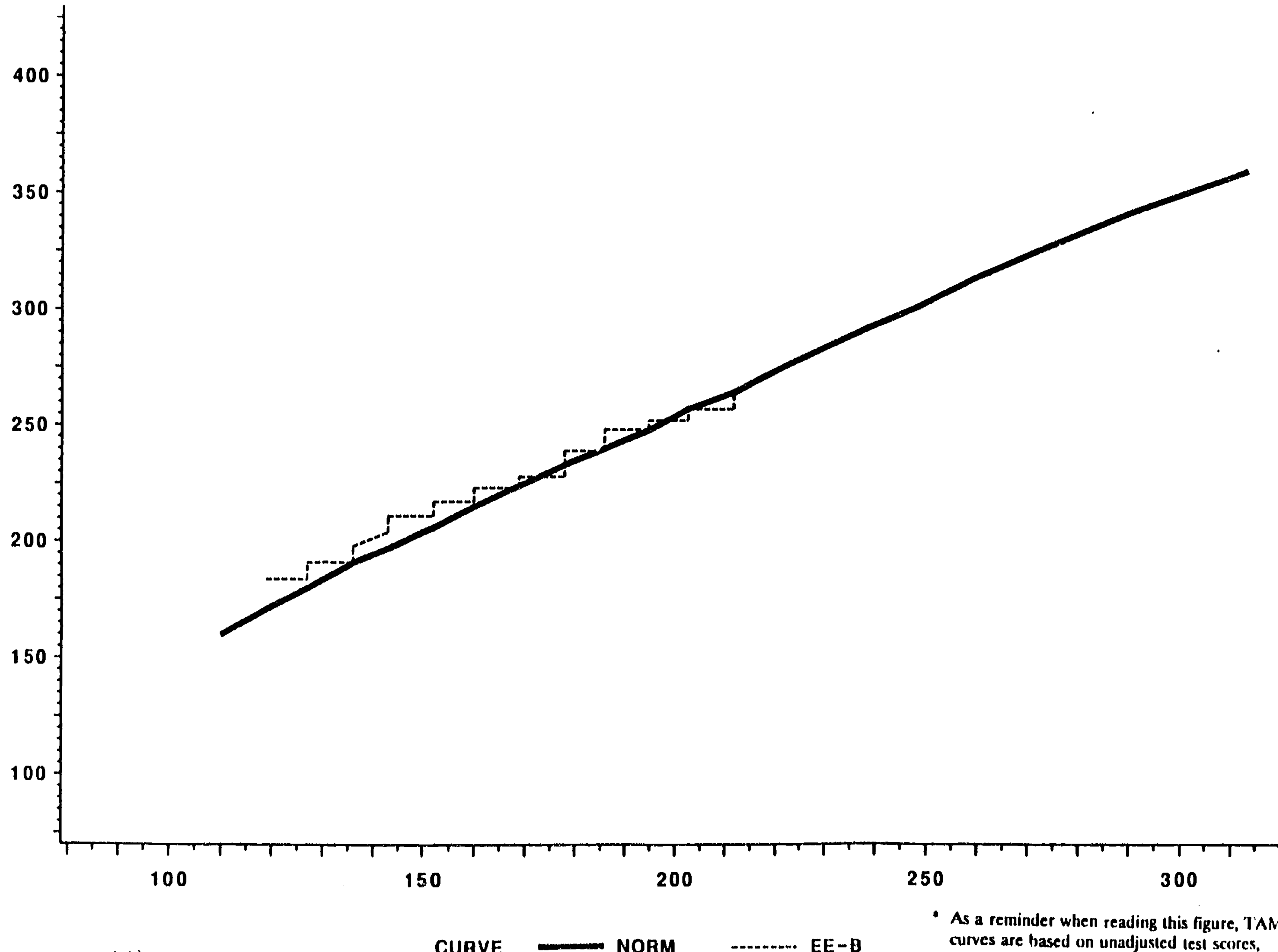
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 160

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District B

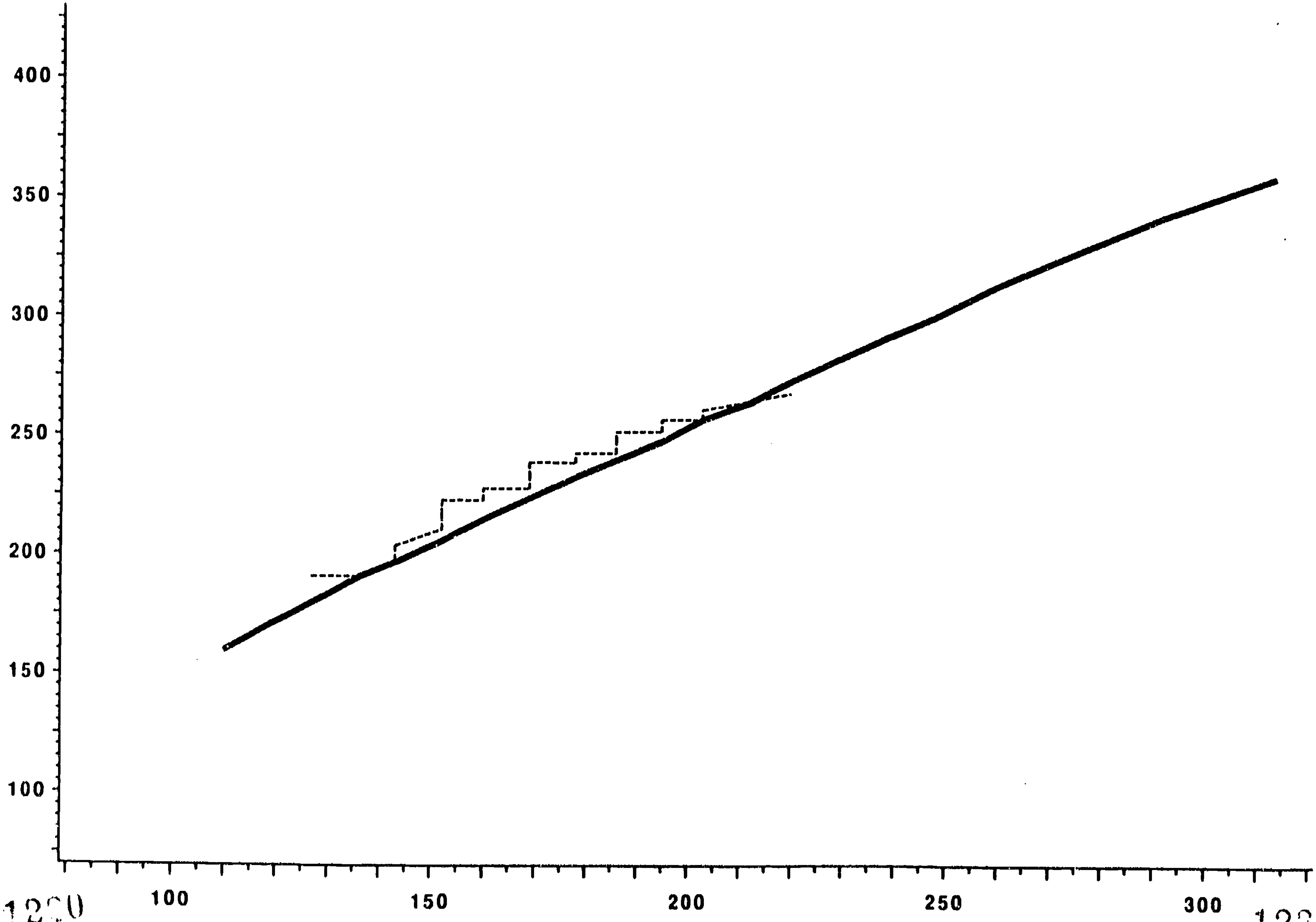


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 161

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District C



CURVE ——— NORM -------- EE-C

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual

532

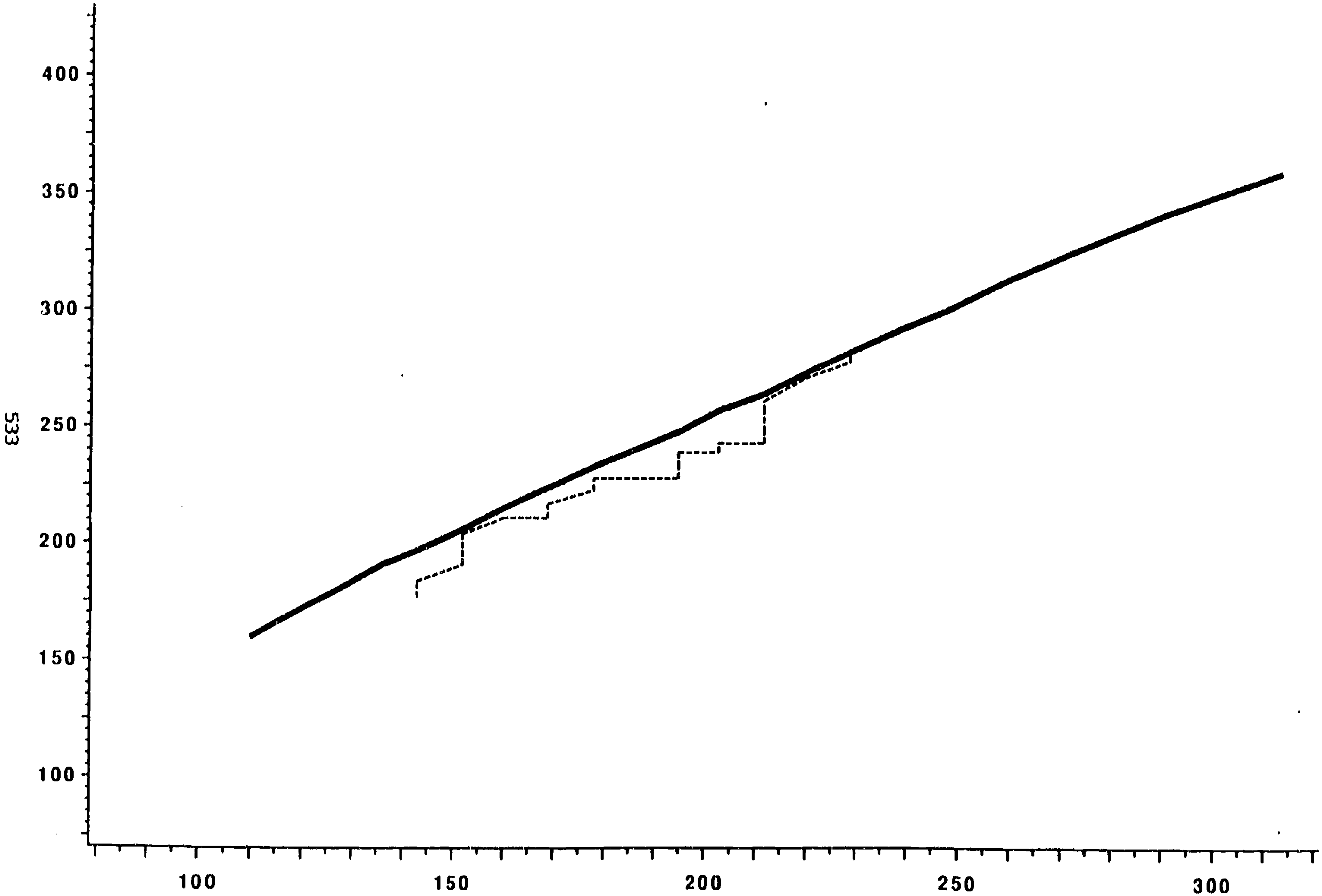
1220

1221

Figure 162

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program in District H



CURVE ——— NORM - - - - - - - - EE-H

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

4. <u>First Grade to Third Grade</u>:

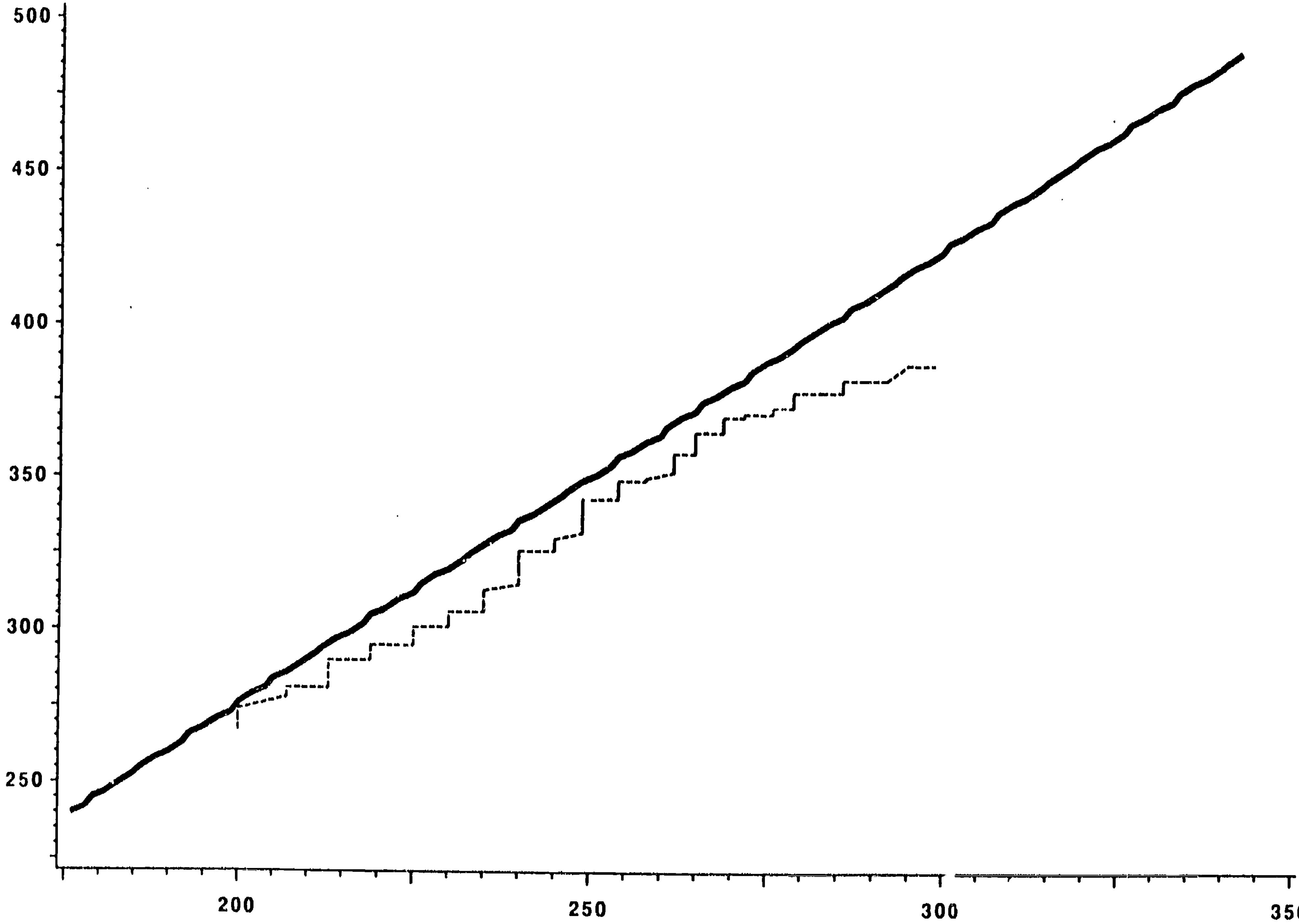
A.	**Grade Span:**	**First Grade**	**to**	**Third Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

Figure 163 suggests that immersion strategy students as a group lost ground in mathematics relative to this norming population, with those students who had higher first grade skills possibly losing more ground than those with lower first grade skills. Early-exit students with lower first grade skills appear to have grown at nearly the same rate relative to this norming population (see Figure 164). However, as with the immersion strategy students, those early-exit students with higher first grade skills lost ground relative to this norming population. The growth rate for late-exit students was similar to that of immersion strategy and early-exit students: a slower growth rate in mathematics relative to this norming population (see Figure 165). Nonetheless, those late-exit students with average or higher first grade entry scores relative to this norming population appear to have increased their mathematics skills almost at the same rate as this norming population. These results are not consistent with the predictions made for each of the instructional models. It was expected that as mathematics was assessed in English, immersion strategy and early-exit students would do as well as this norming population. This did not occur. Both immersion strategy and early-exit students seemed to grow more slowly than this norming population. It was predicted that late-exit students would not learn as fast as this norming population or the immersion strategy or early-exit students when tested in English. This also did not occur. Late-exit students seemed to grow as fast as immersion strategy and early-exit students, given their scores at spring of first grade. This finding presumably reflects the facilitative effects of primary language instruction.

Figure 163

Math Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



535

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

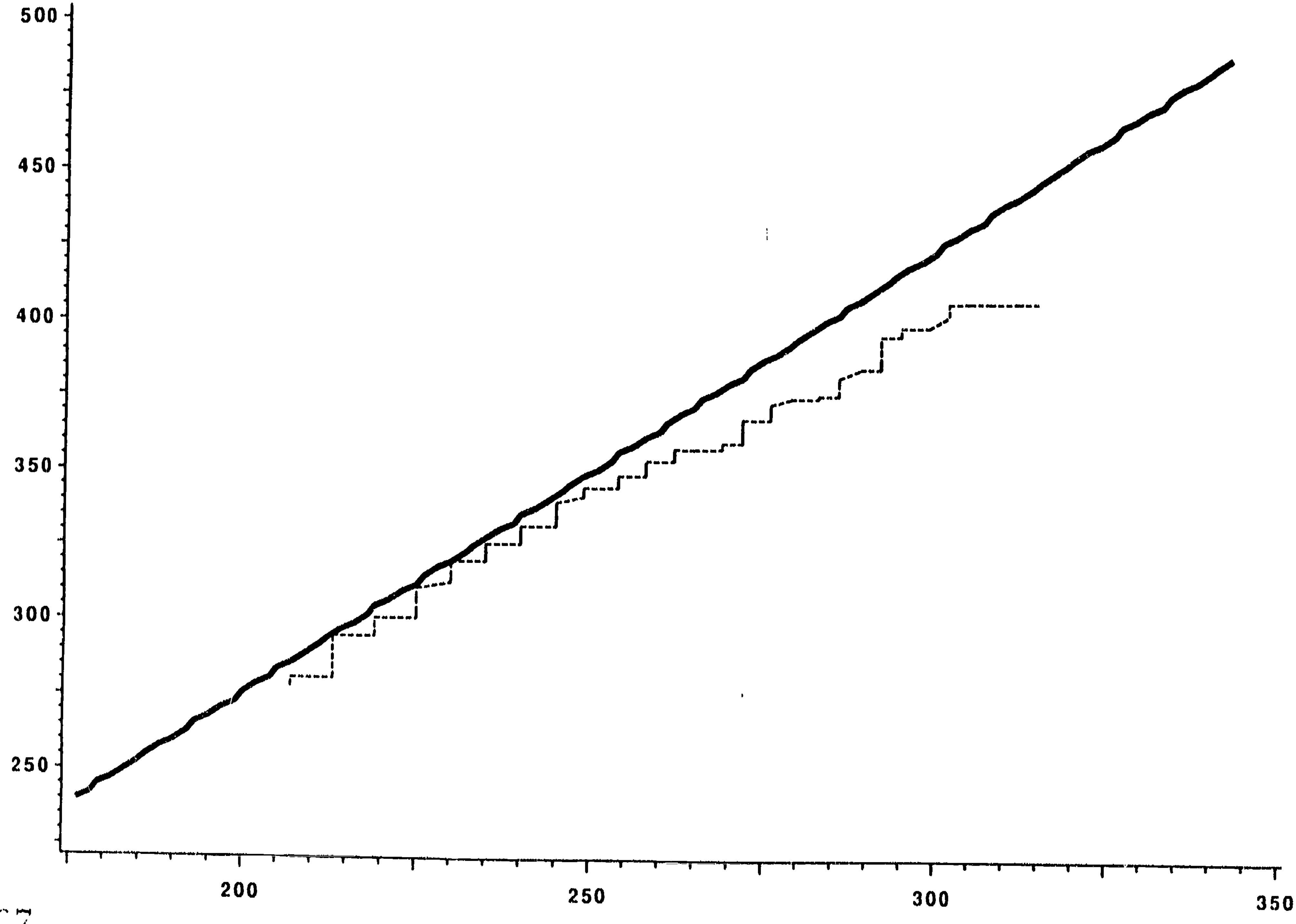
ERIC

Figure 164

Math Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

536

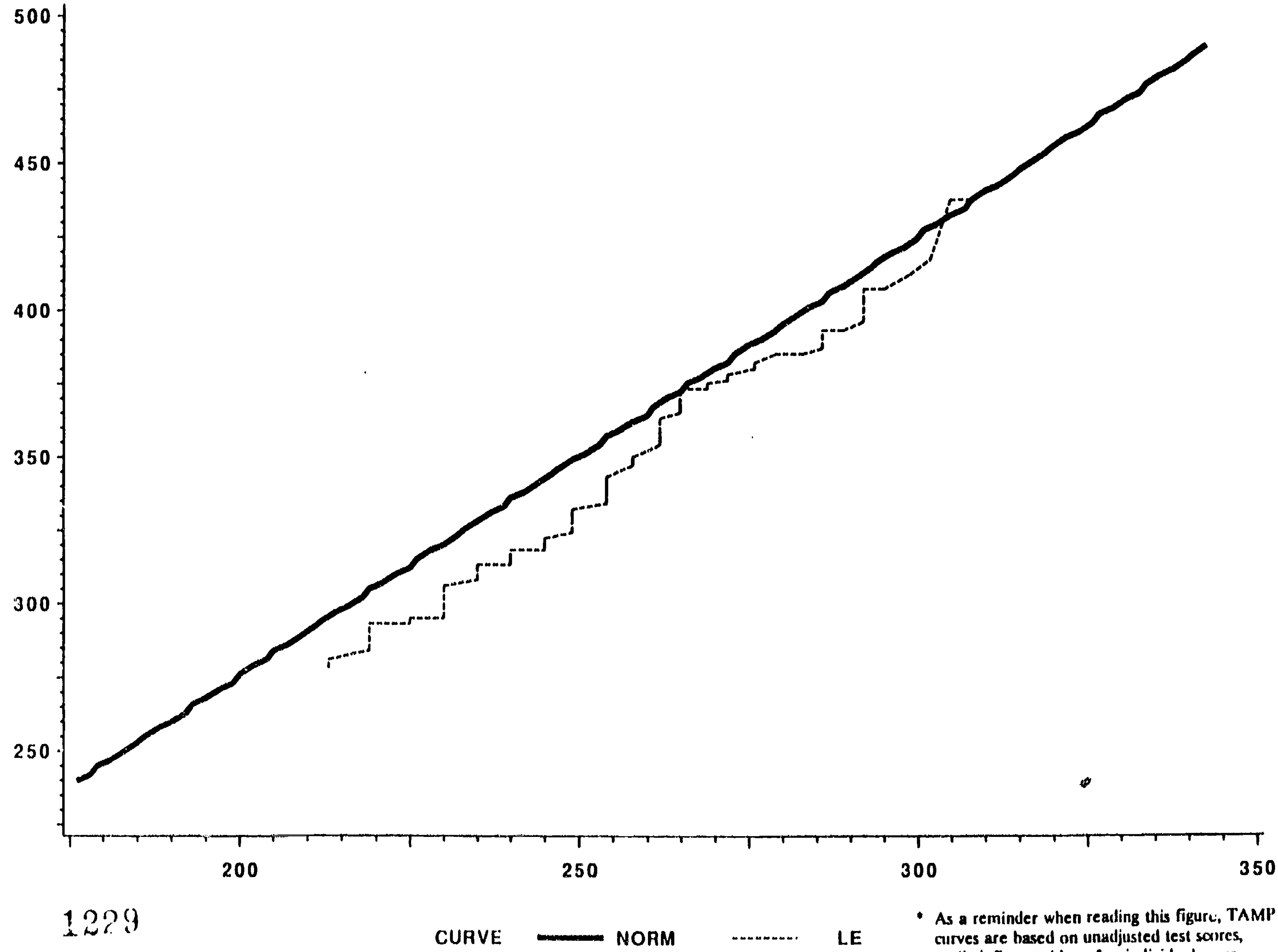


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 165

Math Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

B.	**Grade Span:**	**First Grade**	to	**Third Grade**
	Test Date:	**Spring**	to	**Spring**
	Language:	**English**	to	**English**
	Content:	**Language**	to	**Language**

As in the comparison of spring kindergarten to spring first grade English language skills, immersion strategy and early-exit students increased their English language skills at the same rate relative to this norming population (see Figures 166 and 167). Immersion strategy students with higher English language skills in first grade tended to learn at a slightly faster rate relative to this norming population. Those immersion strategy students with lower initial English language scores seemed to grow at a rate somewhat lower than that of this norming population. In other words, the immersion strategy students appear to show an increasing range of scores: the high scores get relatively higher and the low scores get relatively lower. For early-exit students, overall growth is about the same as this norming population. Similarly, late-exit students with lower and average initial English language scores seemed to learn at the same rate as this norming population (Figure 168). However, those students with higher first grade English language skills appear to have been losing ground slightly relative to this norming population. While this loss was predicted from the late-exit instructional model, it was not predicted that students who had average and lower English language skills at spring of first grade would do as well as this norming population, even though they had limited instruction in English. Once again, perhaps the English language skills represented by these lower levels are sufficiently basic that students could develop them from exposure to English in their environment, and/or these skills were adequately covered in the initial English language instruction received.

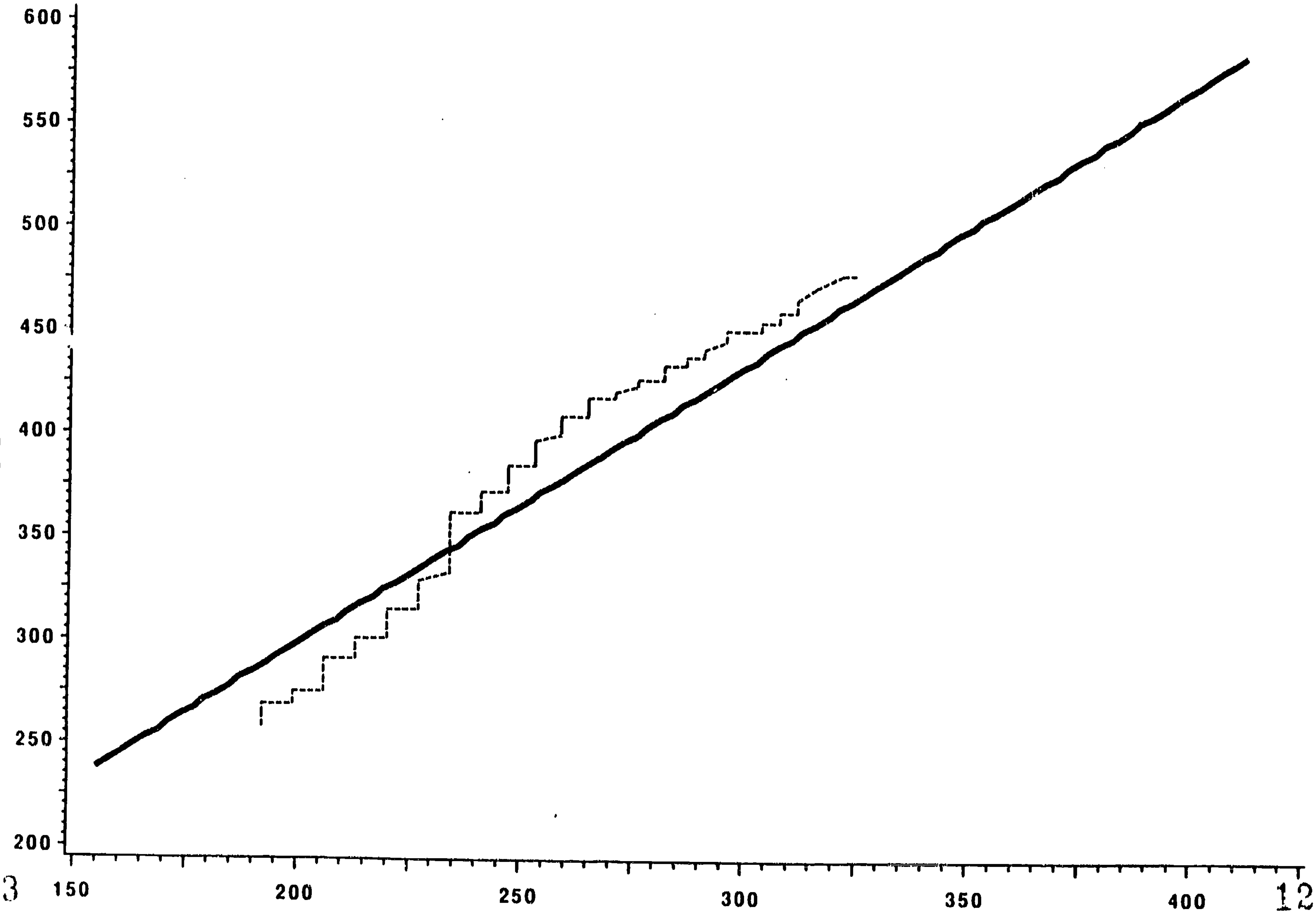
In sum, students in all three programs seemed to post growth rates in English language skills more or less comparable to those of this norming population. While this was predicted for immersion strategy and early-exit students, it was not predicted for late-exit students. The latter finding might be interpreted as supporting the notion that continued development of their primary language would make it easier for these students to acquire English language skills from their environment and/or from their limited instruction in English.

ERIC

Figure 166

Language Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



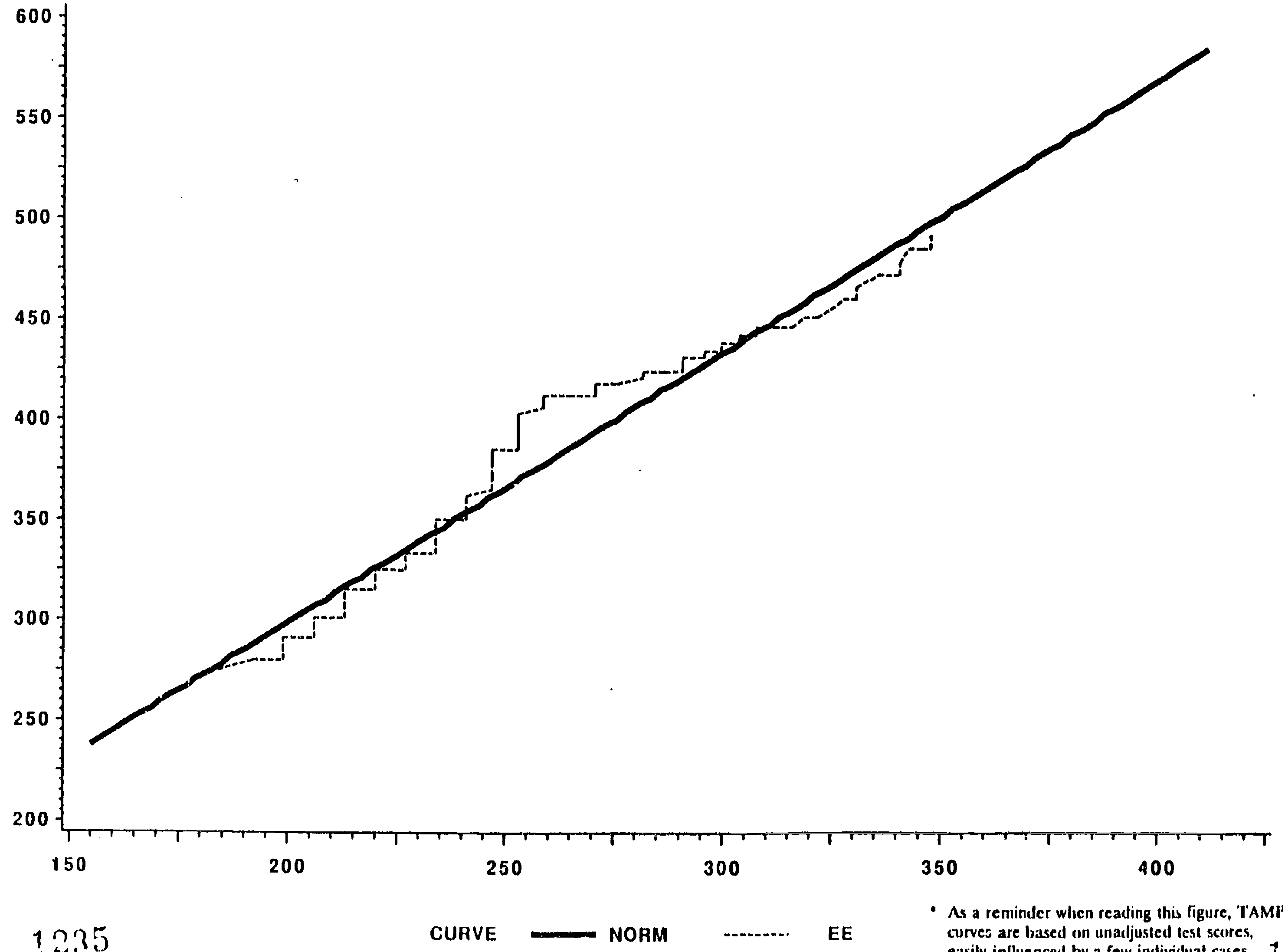
CURVE ▬▬▬ NORM ┅┅┅ IS

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 167

Language Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

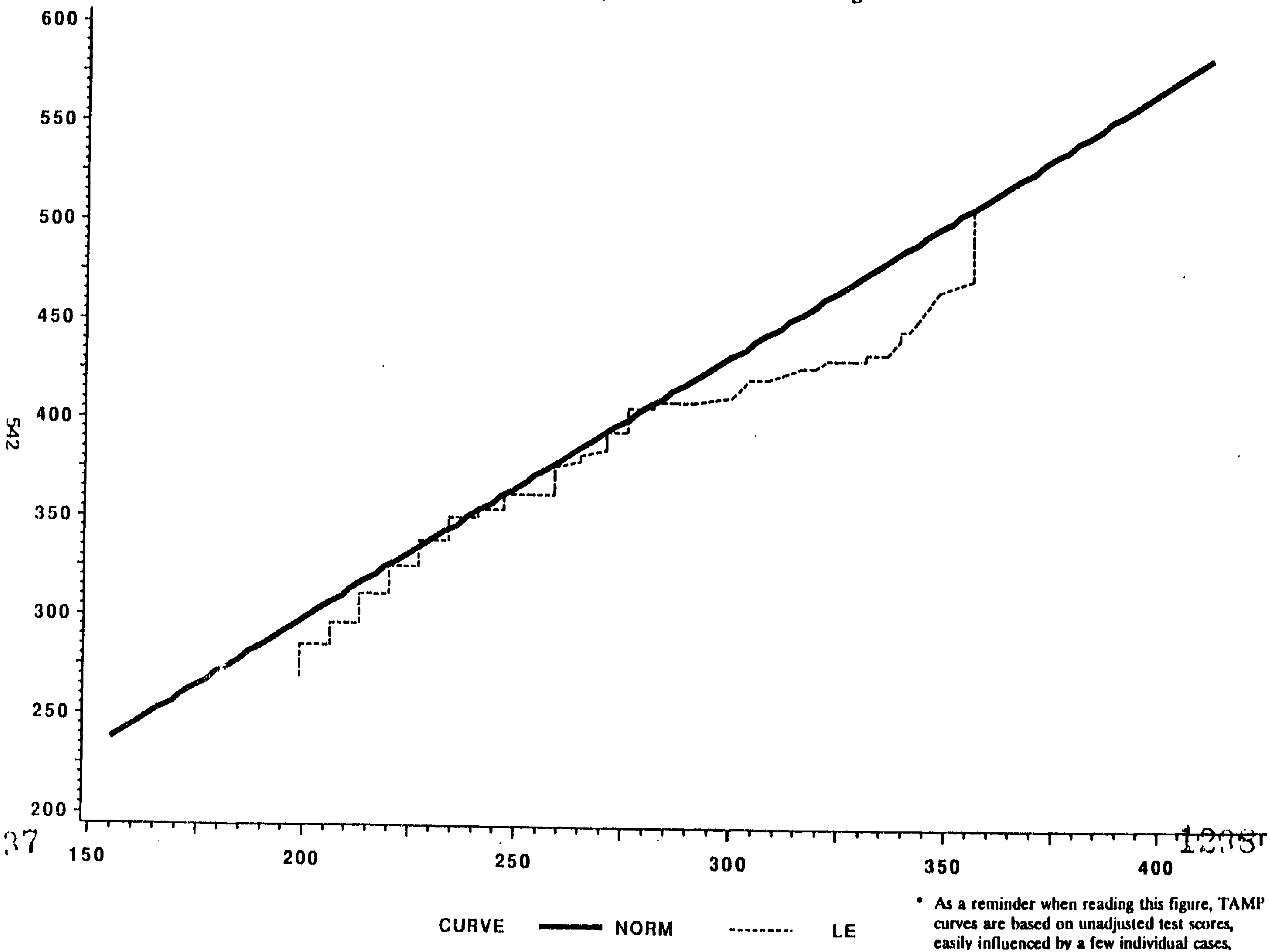


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 168

Language Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

C. **Grade Span: First Grade to Third Grade**
Test Date: Spring to Spring
Language: English to English
Content: Reading to Reading

Grade Span:	**First Grade**	**to**	**Third Grade**
Test Date:	**Spring**	**to**	**Spring**
Language:	**English**	**to**	**English**
Content:	**Reading**	**to**	**Reading**

Immersion strategy and early-exit students appear to have almost the same growth in English reading skills from first to third grade relative to this norming population (see Figures 169 and 170). Immersion strategy students seemed to learn to read slightly faster than this norming population given their status at spring of first grade. Early-exit students appeared to grow as fast as this norming population. Interestingly, given their lower growth from fall kindergarten to spring first grade (see Figure 154), late-exit students also appeared to learn to read in English at nearly the same rate from spring first grade to spring third grade as this norming population, except at the uppermost end of the distribution (see Figure 171). It seems that the predicted growth of immersion strategy students was realized. Immersion strategy students appeared to learn English reading skills slightly faster than this norming population. Early-exit students seemed to exhibit approximately the same growth rate as this norming population, with the slight exception that early-exit students with the highest first grade reading skills might be losing ground relative to this norming population. Late-exit students with higher first grade reading scores seemed to be learning slower than this norming population. Late-exit students with average or below-average first grade reading skills appear to learn at about the same rate as this norming population.

Once again, for all three programs, the finding that students' overall performance seemed to be comparable to, and in some instances better than, this norming population from first grade to third grade attests to the merits of these special support services. While it was predicted that immersion strategy and early-exit students would demonstrate improved English reading skills by virtue of the greater English instruction in these two programs, the finding that late-exit students shared almost comparable growth from first grade to third grade relative

ERIC

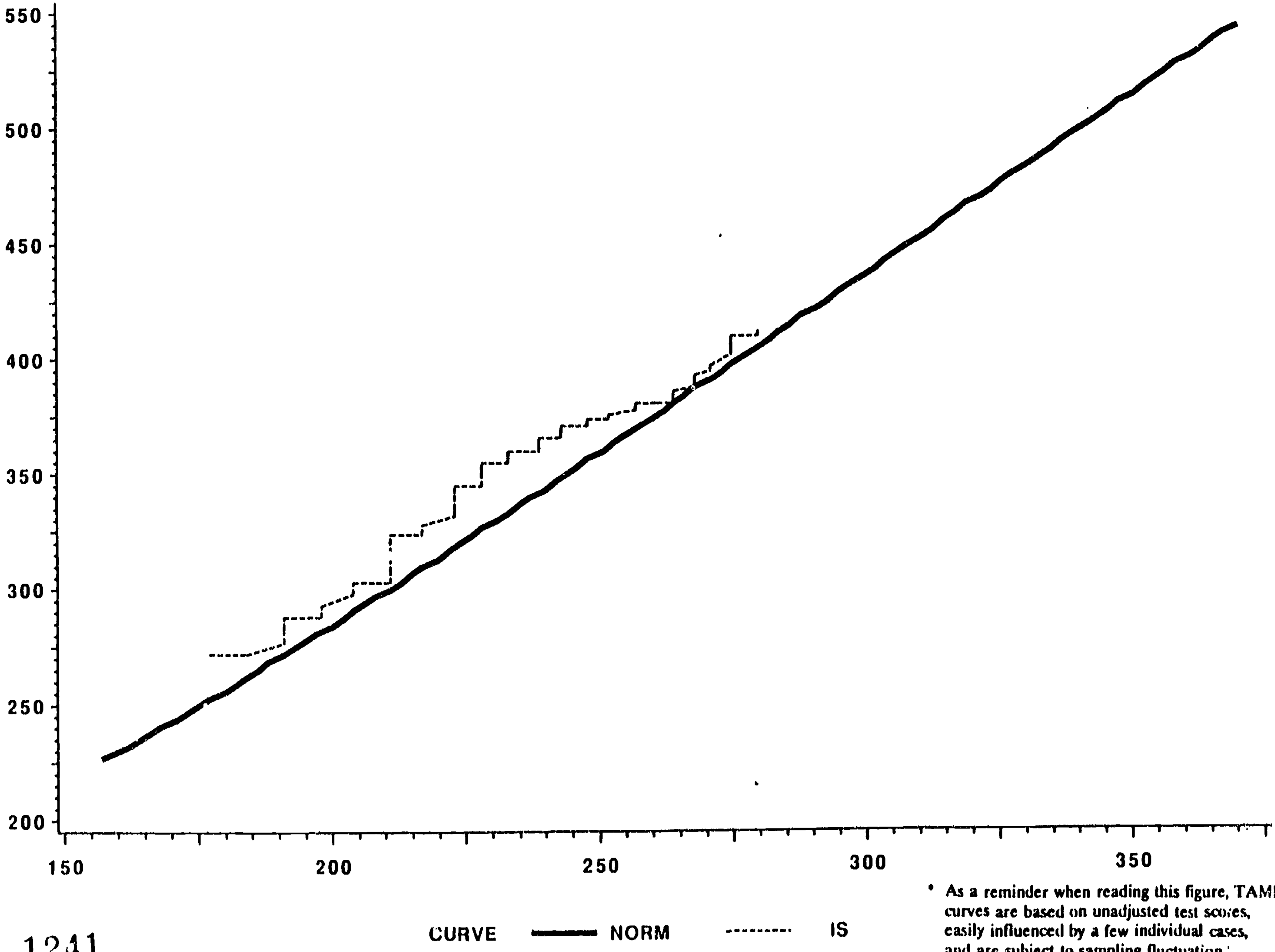
to this norming population for students with average or below-average spring first grade scores was not predicted by virtue of the lower amount of English instruction provided. The latter outcome appears to be consistent with the notion that continued development of primary language skills facilitates the acquisition of second language skills.

ERIC

Figure 169

Reading Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Immersion Strategy Program



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

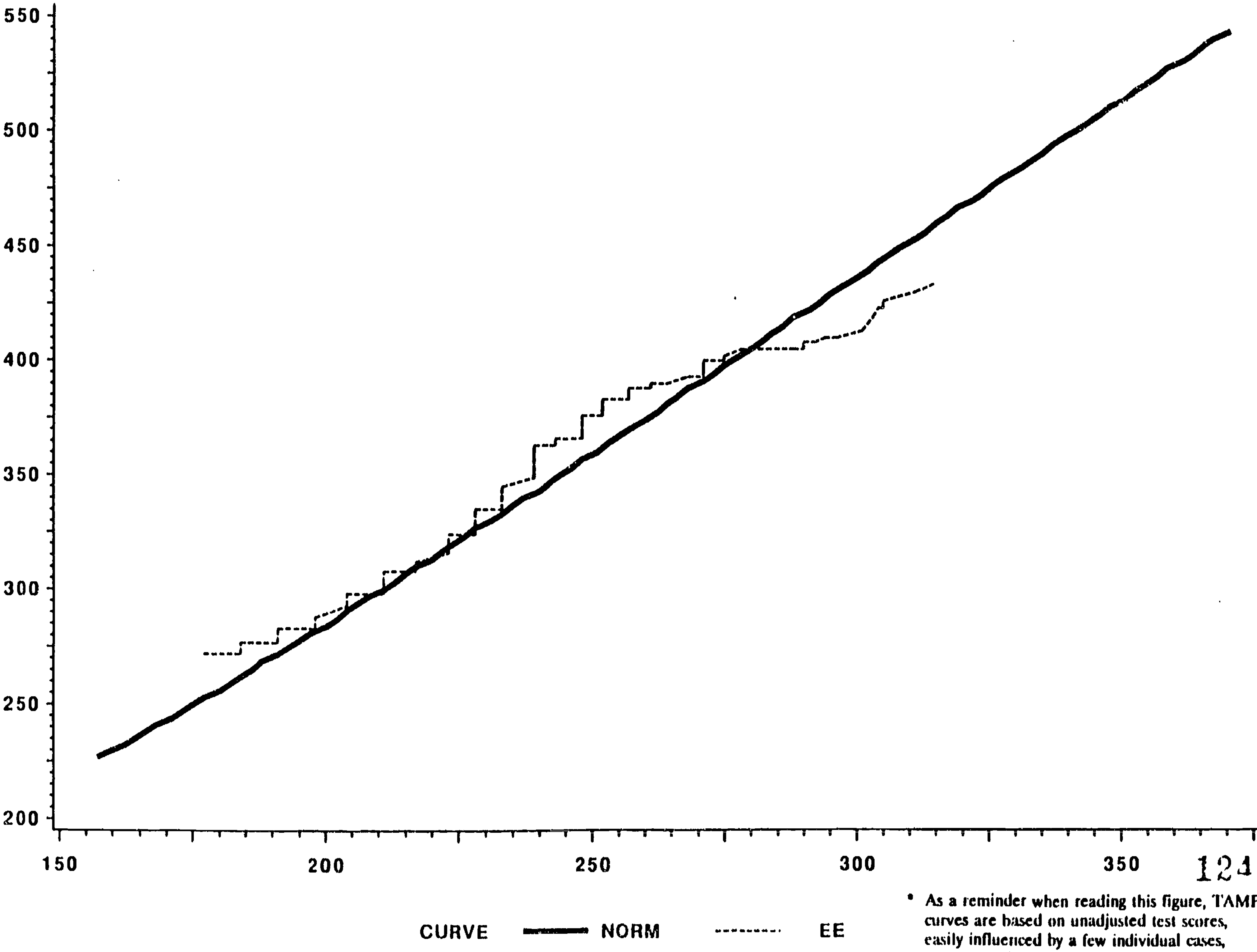
545

ERIC

Figure 170

Reading Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Early-Exit Program

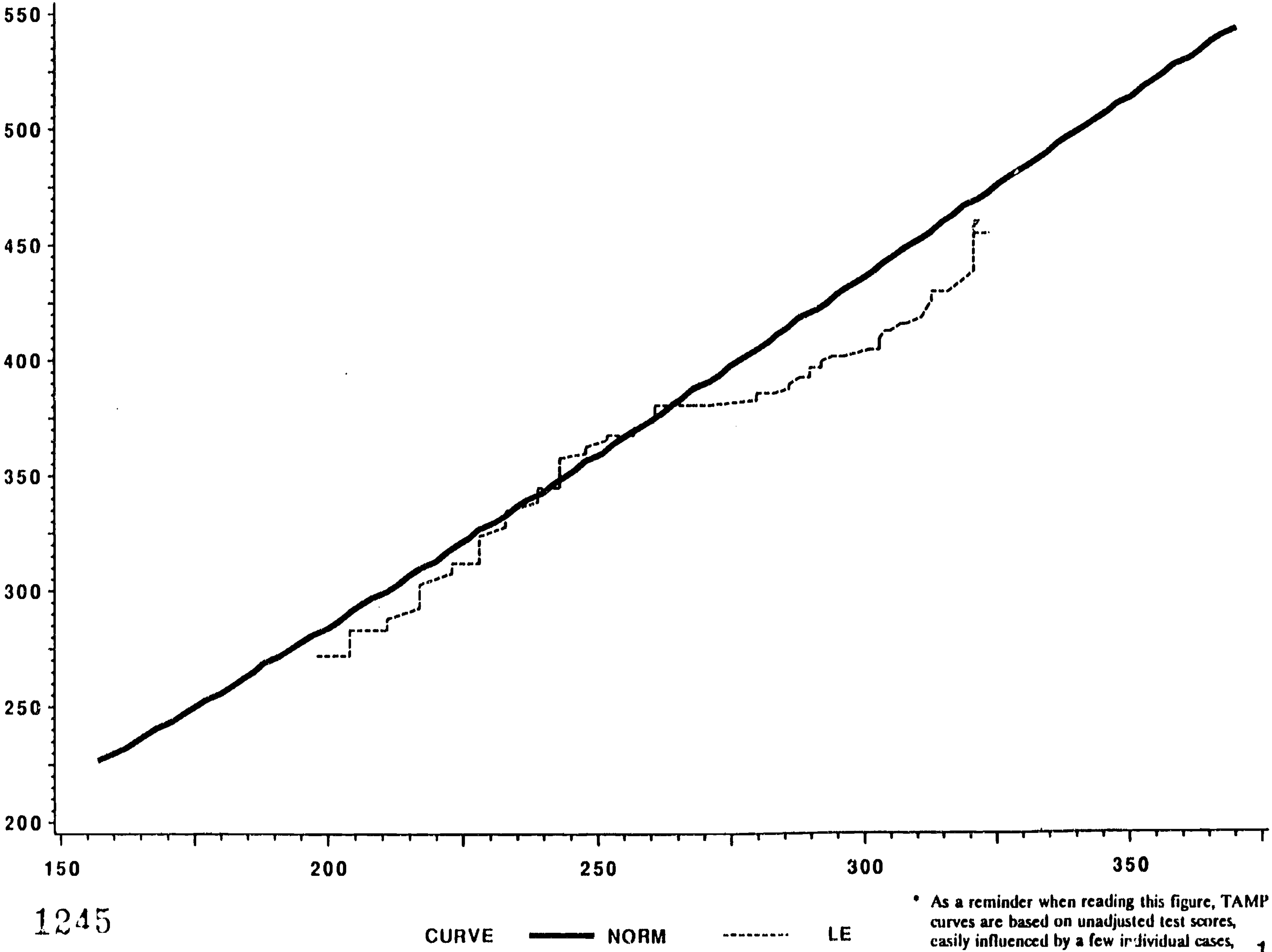


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 171

Reading Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



547

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Summary

In sum, the TAMP analyses relating the growth of immersion strategy, early-exit, and late-exit students to this norming population, with some exceptions, reveal a great deal of consistency with predictions made of each program from their respective instructional models and, as a whole, are fairly consistent with the findings presented in Chapters III through V (see Table 262). It was predicted and it was found that the growth in mathematics, language, and reading for immersion strategy and early-exit students is comparable to one another relative to this norming population. And, in general, these students were learning at or almost at the same rate as this norming population. It was also seen that the initial upward jog in academic growth rate that was noted for immersion strategy students in Chapters III and IV occurs during the first grade year, and that it dissipates by the end of third grade. That is, immersion strategy and early-exit students appear to end kindergarten with approximately the same achievement level, but they end first grade with immersion strategy students having slightly higher achievement than early-exit students. However, both early-exit and immersion strategy students end the third grade with roughly the same achievement level. In a sense, it appears that the extra instruction provided in English to immersion strategy students provides them with an upward jog in their growth rate. However, this growth is short-lived, affecting their achievement for only one year. Thereafter, their rate of growth appears to be similar to that of students in the early-exit program.

Noteworthy is that in spite of less instruction in English, late-exit students appear to be learning as fast or almost as fast as this norming population in mathematics, English language, and English reading. This finding was unexpected. It is proposed that this results from their increased proficiency in their primary language which facilitates their learning of English from their environment and from their limited instruction in English within their instructional program.

Table 262

Summary of TAMP Analyses Results:
Comparison of the Growth Rate for Each Study Program
Relative to the Growth Rate of This Norming Population

		Immersion Strategy	Early-Exit	Late-Exit
1.	Kindergarten/Kindergarten (Fall/Spring)			
A.	Span-Mathematics/Span-Mathematics	-	=	=
B.	Span-Mathematics/Eng-Mathematics	=/-	=	=/-
C.	Span-Language/Span-Language	-	=/-	-
D.	Span-Language/Eng-Language	=/-	=/-	-
2.	Kindergarten/First Grade (Spring/Spring)			
A.	Span-Mathematics/Eng-Mathematics	+	+	+
B.	Eng-Mathematics/Eng-Mathematics	+	+	+
C.	Span-Language/Eng-Language	+	=	=
D.	Eng-Language/Eng-Language	+	+	+
E.	Span-Language/Eng-Reading	+	+	+
F.	Eng-Language/Eng-Reading	+	+	+
3.	Kindergarten/First Grade (Fall/Spring)			
A.	Span-Mathematics/Eng-Mathematics	+	+	+
B.	Span-Language/Eng-Language	=	=	=
C.	Span-Language/Eng-Reading	=	=	=/-
4.	First Grade/Third Grade (Spring/Spring)			
A.	Eng-Mathematics/Eng-Mathematics	-	-	-
B.	Eng-Language/Eng-Language	=	=	=
C.	Eng-Reading/Eng-Reading	+/=	=	=

=: Growth rate equals/approximates that of this norming population

+: Growth rate is faster than that of this norming population

-: Growth rate is slower than that of this norming population

ERIC

There are noticeable differences in the growth of mathematics skills for students in each program when their growth rates from kindergarten to first grade are compared to their growth rates from first grade to third grade. Students in each program increased their skills in mathematics at a faster rate than this norming population from spring kindergarten to spring first grade. However, from spring first grade to spring third grade, students in each program lost ground relative to this norming population. That is, students in each program gained ground relative to this norming population from kindergarten to first grade, only to fall back from first grade to third grade. This may reflect the nature of the kindergarten and first grade curriculum which tends to utilize manipulatives and other approaches to shelter, or contextualize, language, thereby making it more meaningful for limited-English-proficient students and easier to learn. As the second and third grade curriculum tends to increase its reliance on language as the primary vehicle for instruction and to decrease the use of strategies to contextualize language, learning would be more difficult and slower.

A different pattern emerges when similar comparisons are made for English language and English reading skills. With two exceptions, students in all three programs seemed to increase their English language and reading skills at the same rate as this norming population. One exception is the decreased growth in English reading skills experienced by late-exit students from spring kindergarten to spring first grade. This decrease is predicted from the late-exit model in that given the limited amount of opportunity to formally develop English language skills (receptive skills at the minimum), English reading development would occur more slowly.

The second exception is the growth rate in English reading by immersion strategy students from first grade to third grade which seemed to be faster than this norming population. This increase in English reading skills is predicted from the immersion strategy model. Given the expanded opportunity to develop English language skills and considering their below-average achievement level in spring of first grade, it is not surprising that the rate of growth in English reading skills would be faster than the growth rate of this norming population.

Of value is that the TAMP analyses graphically support the merits of providing special services to limited-English-proficient students. In each program, students appeared to perform better than expected. Typically, the achievement of language-minority students tends to decrease over time relative to this norming population. In all three programs, target students as a group tended to perform better than, as well as, or almost as well as this norming population.

In sum, the analyses in Chapters III and IV concluded that immersion strategy and early-exit students had comparable skills in mathematics, language, and reading at the end of grade three. These analyses also noted transitory boosts in achievement for immersion strategy students in first grade language and second grade reading skills. The TAMP analyses suggest that immersion strategy and early-exit students do not differ in their patterns of growth relative to this norming population. Of interest, and surprisingly, is the finding that the form of growth for late-exit students relative to this norming population is similar to that of immersion strategy and early-exit students. This suggests that the hypothesis that providing limited-English-proficient students with large amounts of primary language instruction would impede their growth of English language skills does not appear to be supported by these data.

Late-Exit Sites Relative to This Norming Population

The late-exit program in this study is comprised of three school districts, sites LE-D, LE-E, and LE-G. Site D is comprised of one school, there are seven schools in site E, and six schools in site G. The TAMP curves are unstable for site D because the analyses are based on only one school and few students.

Due to data collection restrictions imposed by late-exit site G, fall kindergarten test scores are not available for this site. Any analyses of fall kindergarten data are necessarily limited to late-exit sites D and E. As test scores for site G are available beginning with spring kindergarten, site G is included in the discussions of TAMP results in which spring to spring results are examined. Also, as the TOBE does not provide a measure of reading readiness skills, no reading achievement data are provided for any K-1 analyses.

Late-Exit Site Characteristics

As a reminder, site E provided its students with the greatest amount of primary language instruction, ranging from 87.8% in kindergarten to 42.2% in grade five. In grade six the percentage of instruction in Spanish dropped to 24.2%. In contrast, students in site D received approximately 40% of instruction in Spanish from kindergarten through the fifth grade. In grade six the proportion of Spanish instruction also dropped to 26.3%. In site G, only first and second grade students received substantial amounts of instruction in Spanish (47.3% and 40.1%, respectively). Third grade students received about one-fourth (26.2%) of their instruction in Spanish, and fourth grade students one-sixth (15.6%). Essentially all instruction provided to fifth and sixth graders was in English, with almost no Spanish instruction (7.1% and 4.7%, respectively). Recall that these patterns suggested that while sites D and E more closely fit the late-exit instructional model, site G, while nominally a late-exit program, more closely resembled an early-exit instructional model in the amount of Spanish instruction provided.

Late-exit sites also varied with regard to family, school, and community factors. Site D was comprised of students with slightly higher family income and educational levels and was in a community where Spanish is used daily throughout all private and public sectors, as compared to students in either site E or site G. Site G students came from suburban homes that were essentially low-income and blue collar. Of all of the study sites, site E was in the community with the most needs. Site E has all of the characteristics of an inner city area: high crime, violence, transiency, lack of community resources, minimal school resources, overcrowding, less than ideal school facilities, neighborhoods drastically in need of physical repair, no parks, etc. Without question, site E students have the most stress in their community and school environment. However, this is offset to some degree by the finding that teachers in site E have the most education and training of teachers in all of the study sites.

1. Fall Kindergarten to Spring Kindergarten:

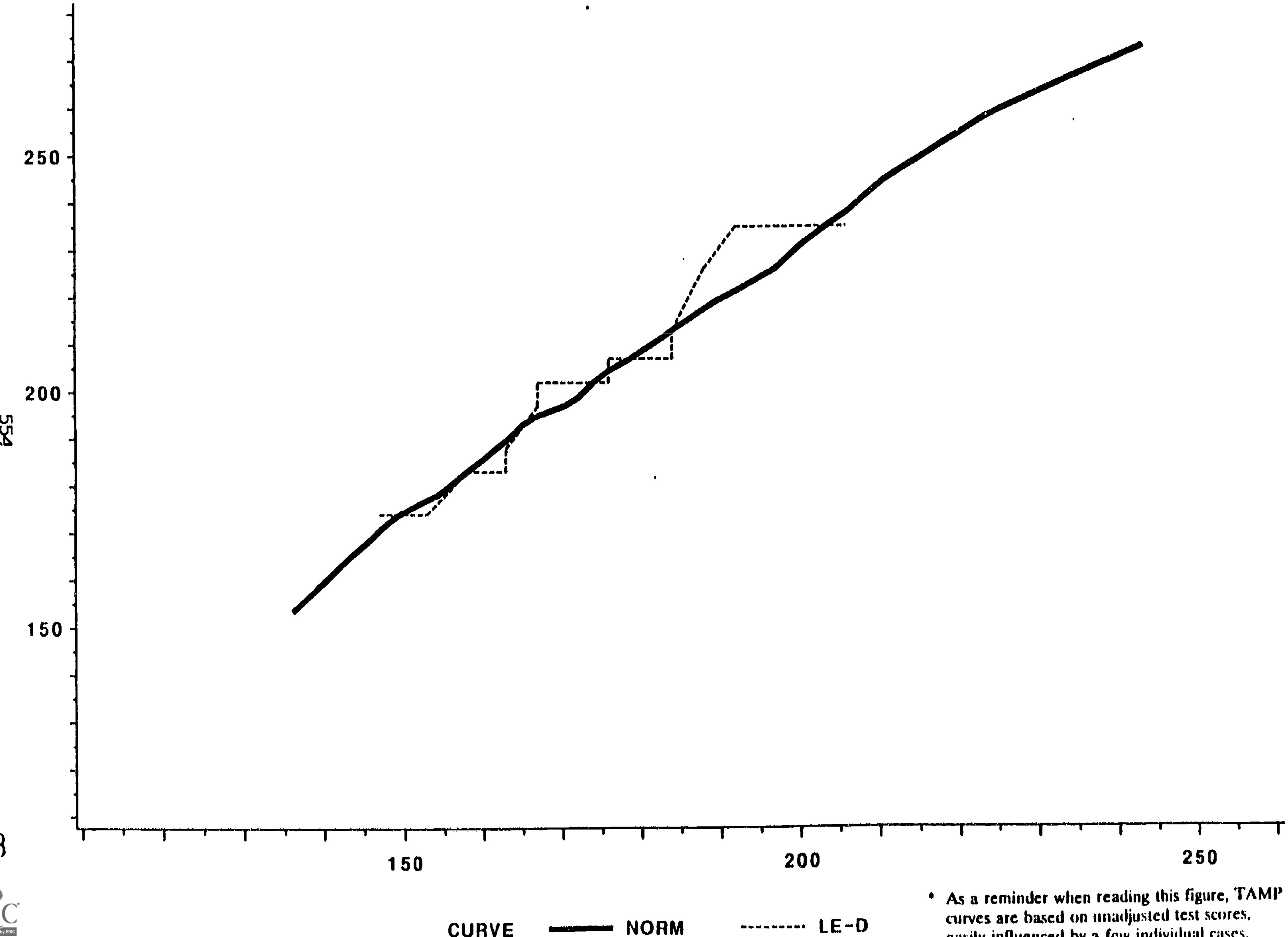
A. Grade Span:	**Kindergarten**	**to**	**Kindergarten**
Test Date:	**Fall**	**to**	**Spring**
Language:	**Spanish**	**to**	**Spanish**
Content:	**Mathematics**	**to**	**Mathematics**

From Figures 172 and 173 it appears that students in site D began kindergarten with slightly higher skills in mathematics than students in site E. Both site D and site E students seemed to increase their skills in mathematics at the same rate as this norming population. These trends are consistent with the predictions made for this instructional model, in that instruction in the primary language would allow for normal progress in the content areas.

Figure 172

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



554

1253

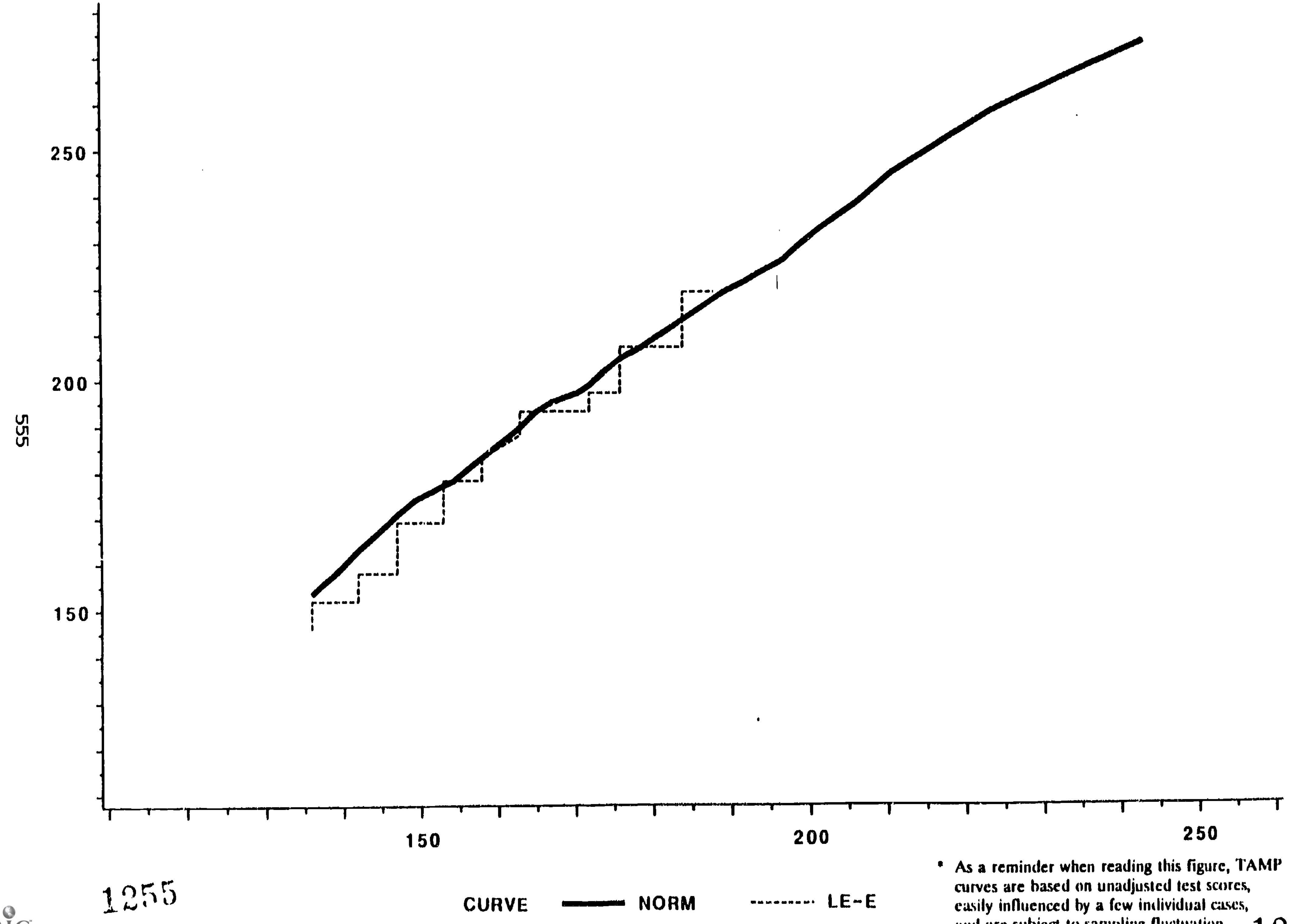
1254

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 173

Spanish Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

B.	**Grade Span:**	**Kindergarten**	**to**	**Kindergarten**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

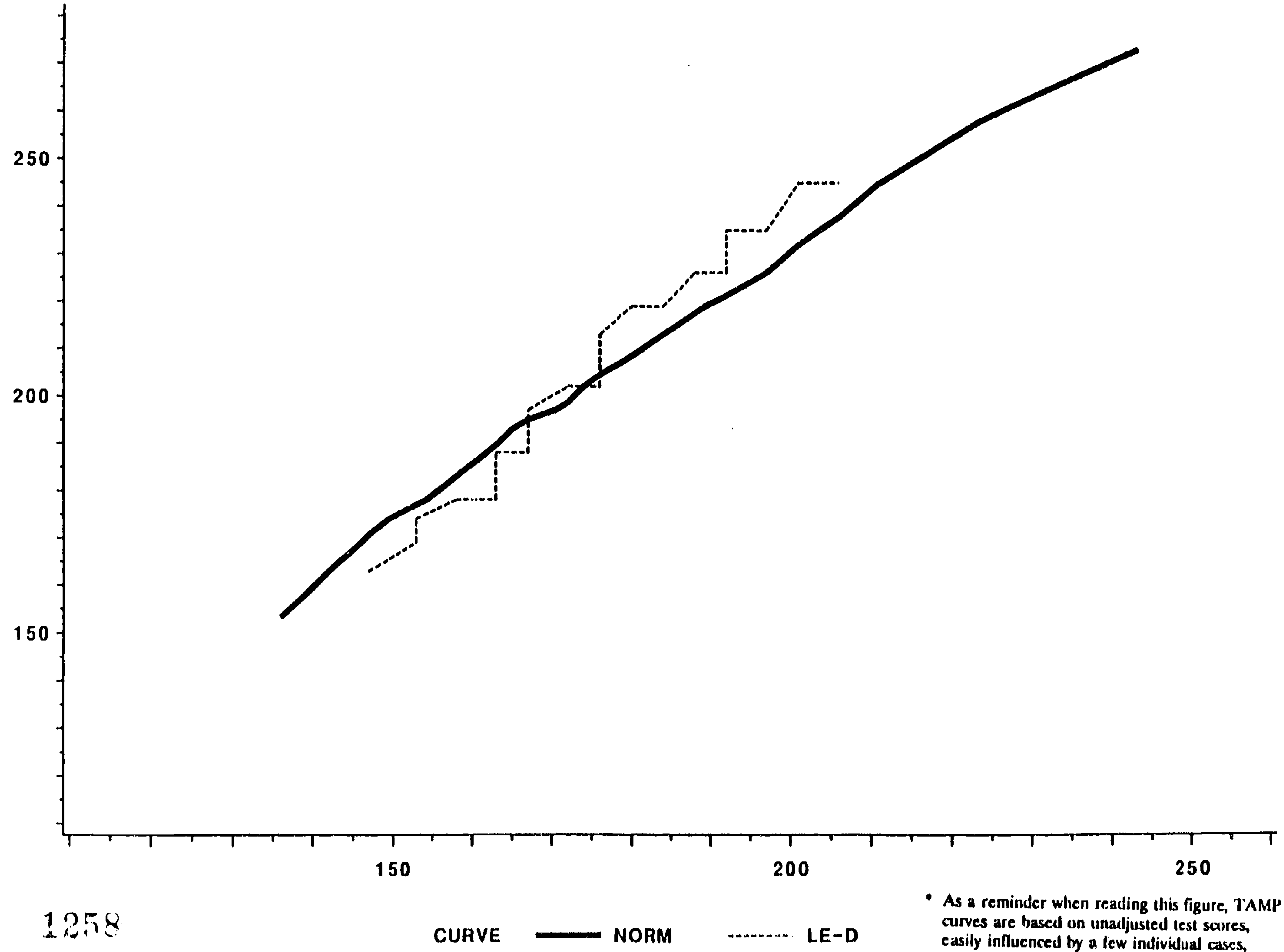
Figures 174 and 175 suggest that students in late-exit site D had higher entry-level mathematics skills than students in site E. This difference presumably reflects the higher socioeconomic level of site D students than that of site E students. Moreover, relative to this norming population, site D students with higher entry-level skills tended to exhibit faster growth in mathematics skills, and those with lower entry-level skills somewhat slower growth. Overall, the growth among site D students seems to be comparable to this norming population. Students in site E appear to have lost ground relative to this normirg population.

ERIC

Figure 174

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D

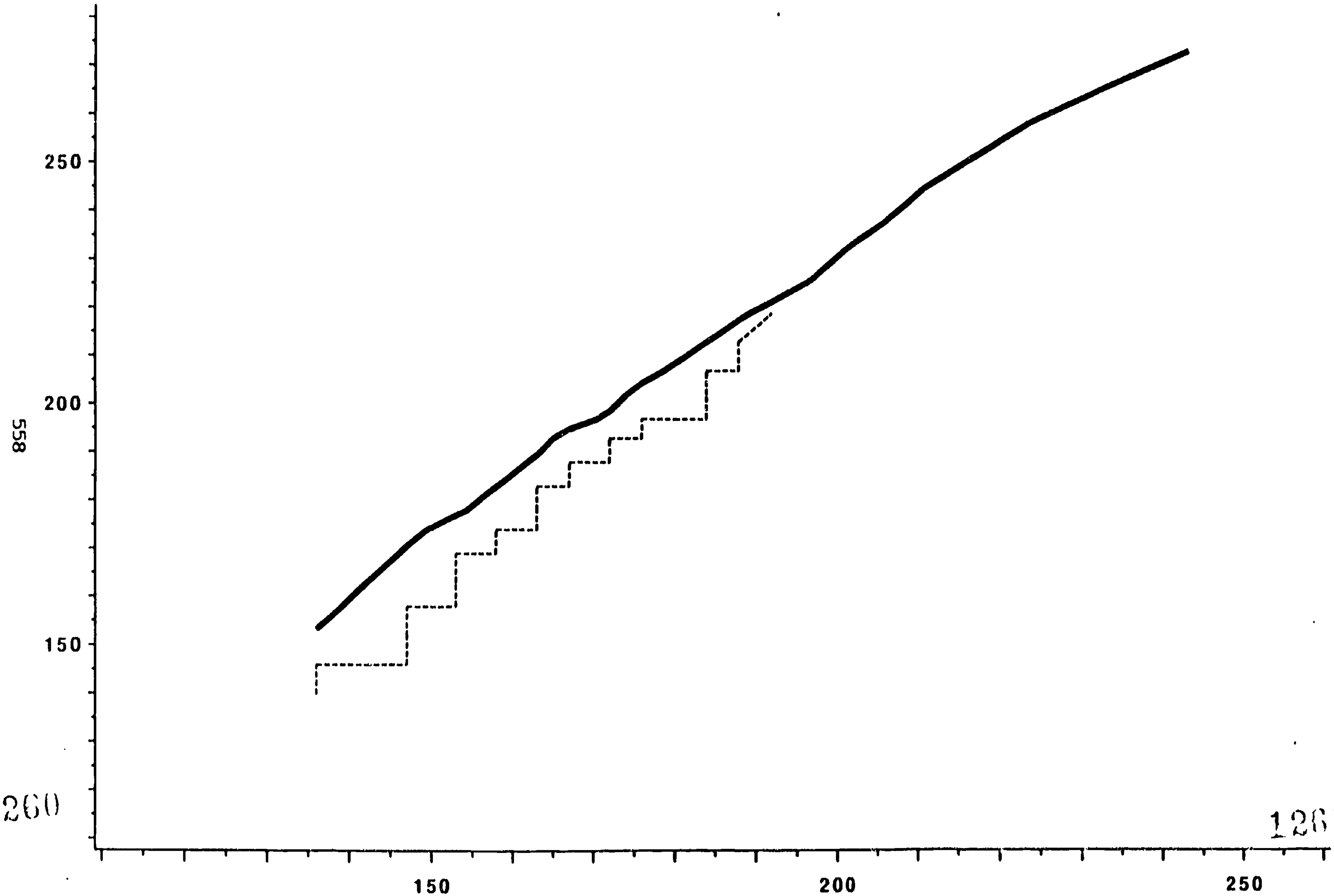


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 175

English Math Spring K vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

C. **Grade Span:**	**Kindergarten**	**to**	**Kindergarten**
Test Date:	**Fall**	**to**	**Spring**
Language:	**Spanish**	**to**	**Spanish**
Content:	**Language**	**to**	**Language**

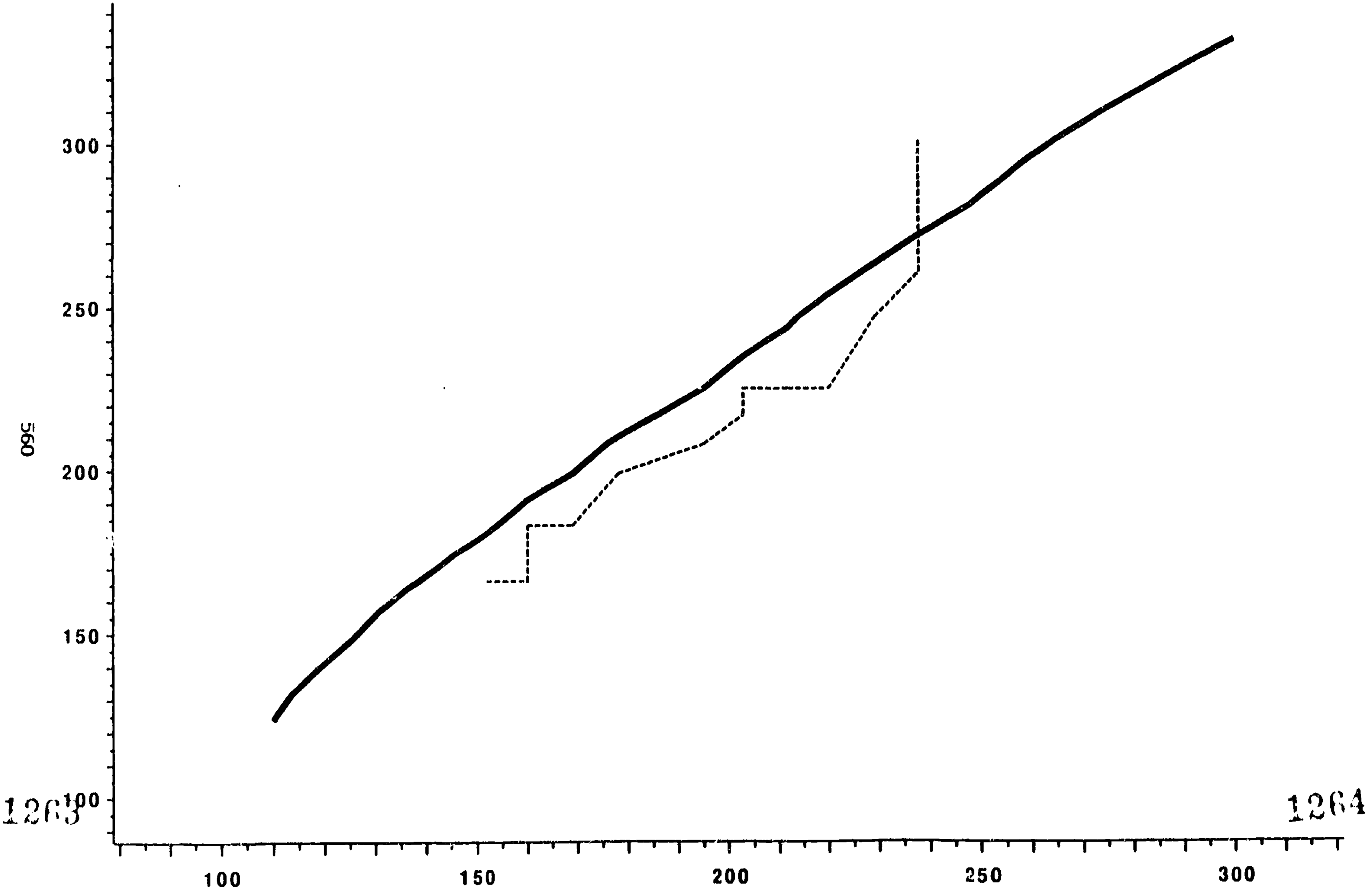
This analysis addresses the issues of how comparable the Spanish language skills of students in both late-exit sites were at the beginning of kindergarten, and of how fast each student group grew in their Spanish language skills during their kindergarten year relative to this norming population. Although the range of initial kindergarten scores is broader for site E students than for site D students, students in both sites began kindergarten with roughly the same Spanish language skills (see Figures 176 and 177). Students in both sites appear to be losing ground compared to this norming population. As noted, the lower range in initial entry-level scores for site E students may reflect the lower socioeconomic and educational levels of the families and community in site E as well as the lower use of Spanish in the larger community as compared to site D.

ERIC

Figure 176

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D

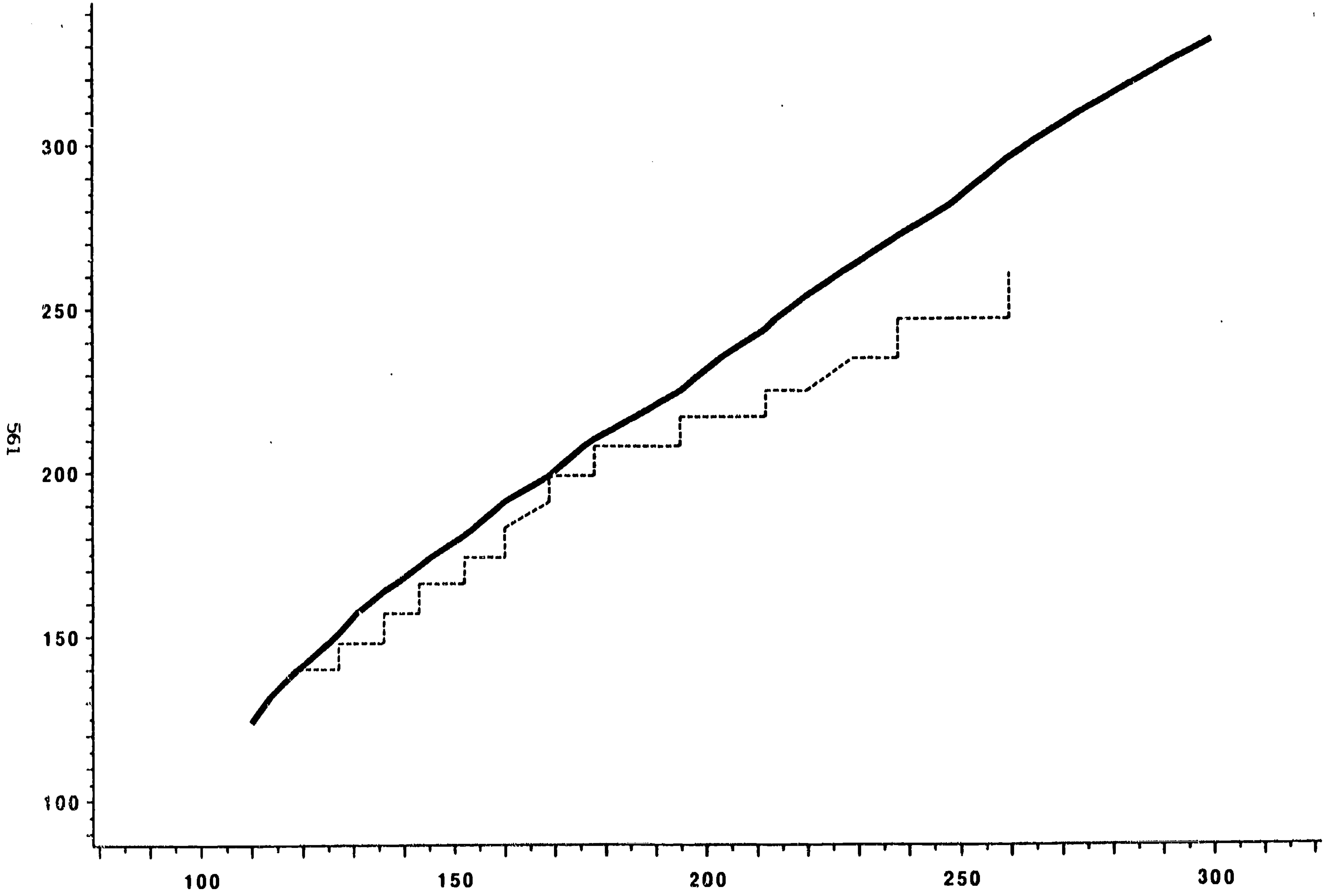


- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

Figure 177

Spanish Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



CURVE ━━━ NORM - - - - - LE-E

- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

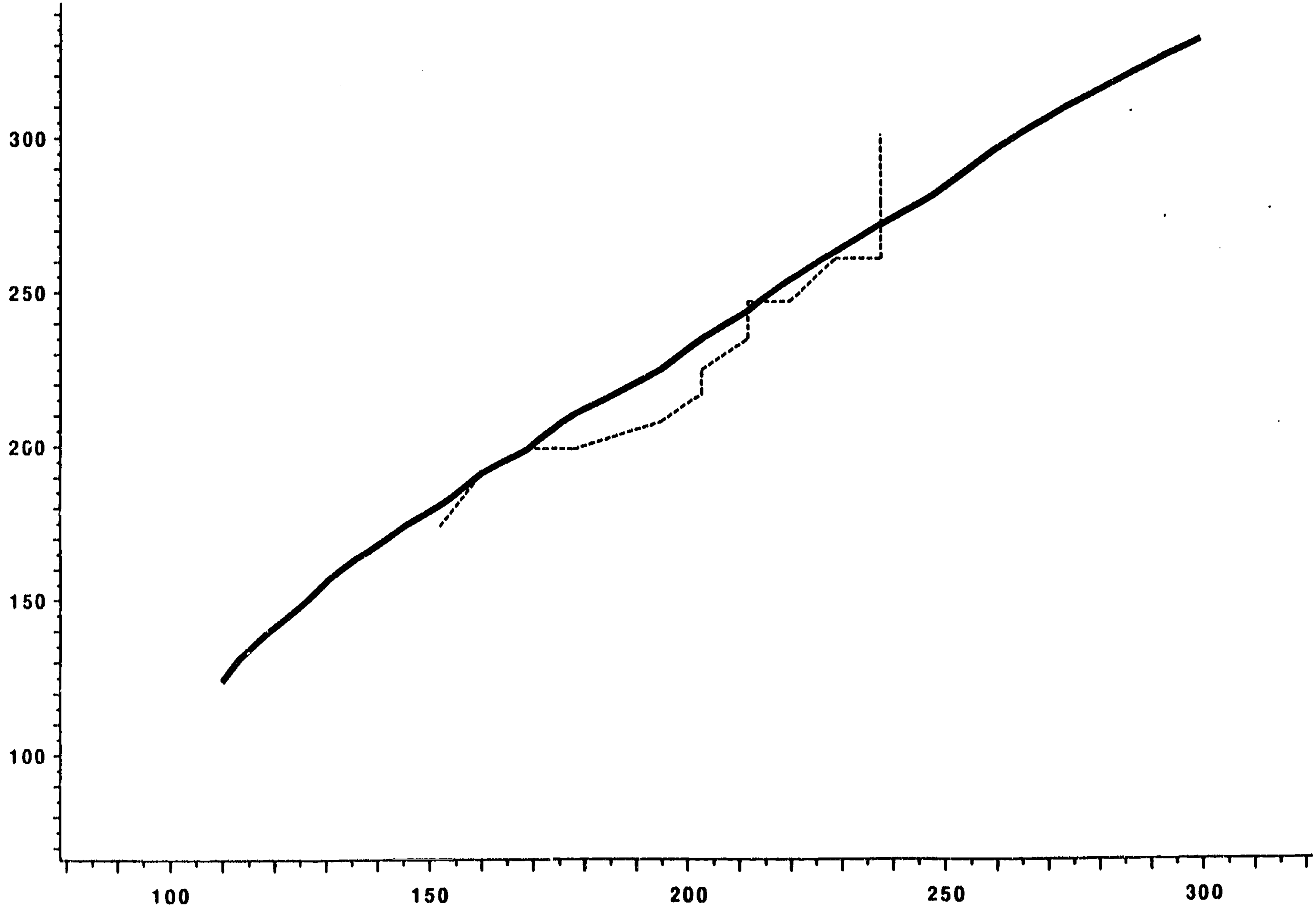
D.	**Grade Span:**	**Kindergarten**	**to**	**Kindergarten**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Language**	**to**	**Language**

Figures 178 and 179 suggest that while students in site E had a lower range of fall kindergarten scores, students in both sites D and E began their kindergarten year with roughly the same level of Spanish language skills. Reflecting the greater amount of English language instruction received, site D students on average grew at approximately the same rate as this norming population and markedly faster than students in site E. This was better than predicted from the late-exit instructional model which suggested that late-exit students should not be progressing at the same rate as this norming population given their limited exposure to English language instruction. Site E students, receiving a more gradual use of English for instruction, appeared to experience a slower growth rate in English language skills relative to this norming population, as expected.

Figure 178

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



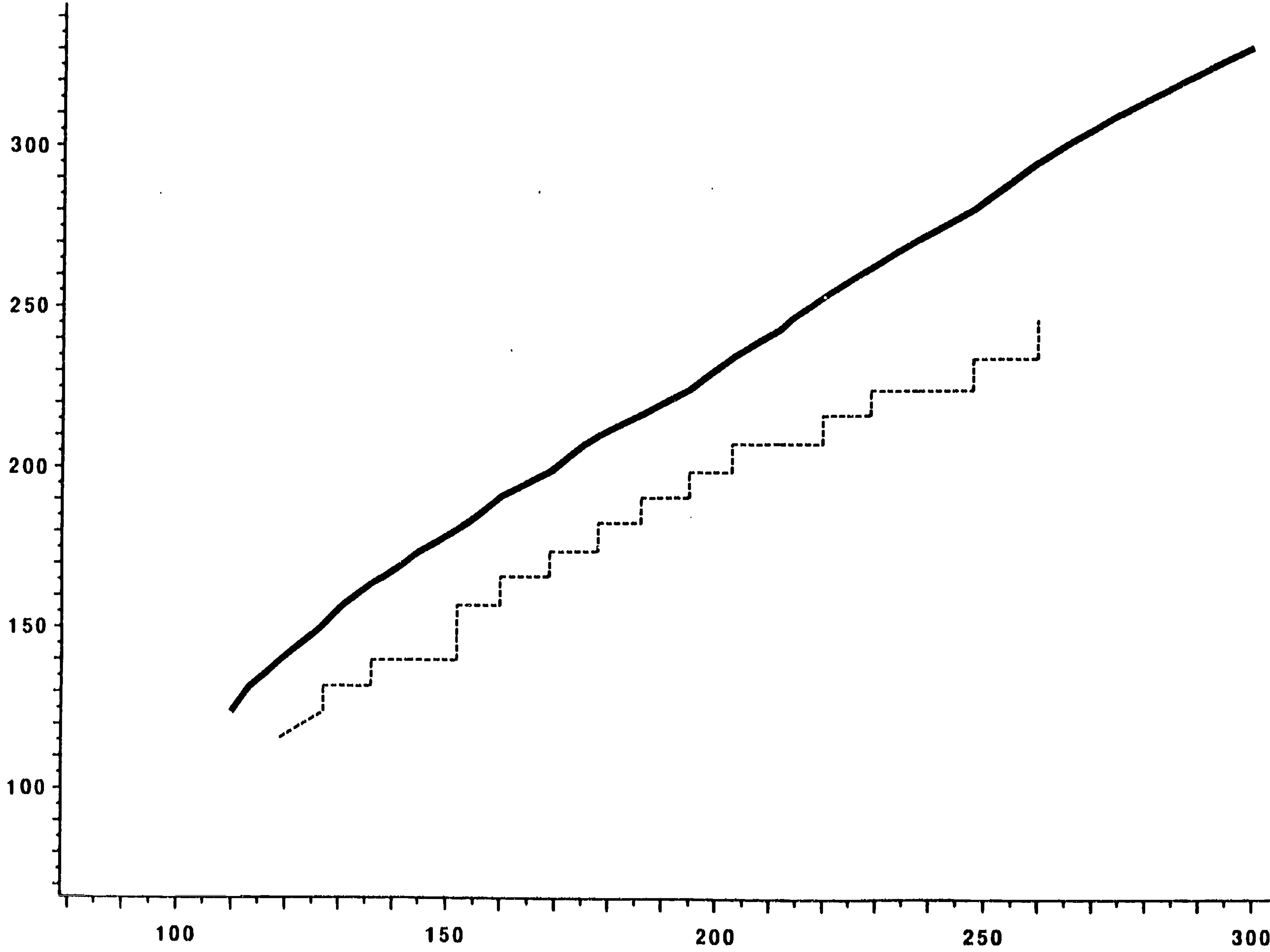
563

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 179

English Language Spring K vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



CURVE NORM LE-E

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC
BEST COPY AVAILABLE

2. Spring Kindergarten to Spring First Grade:

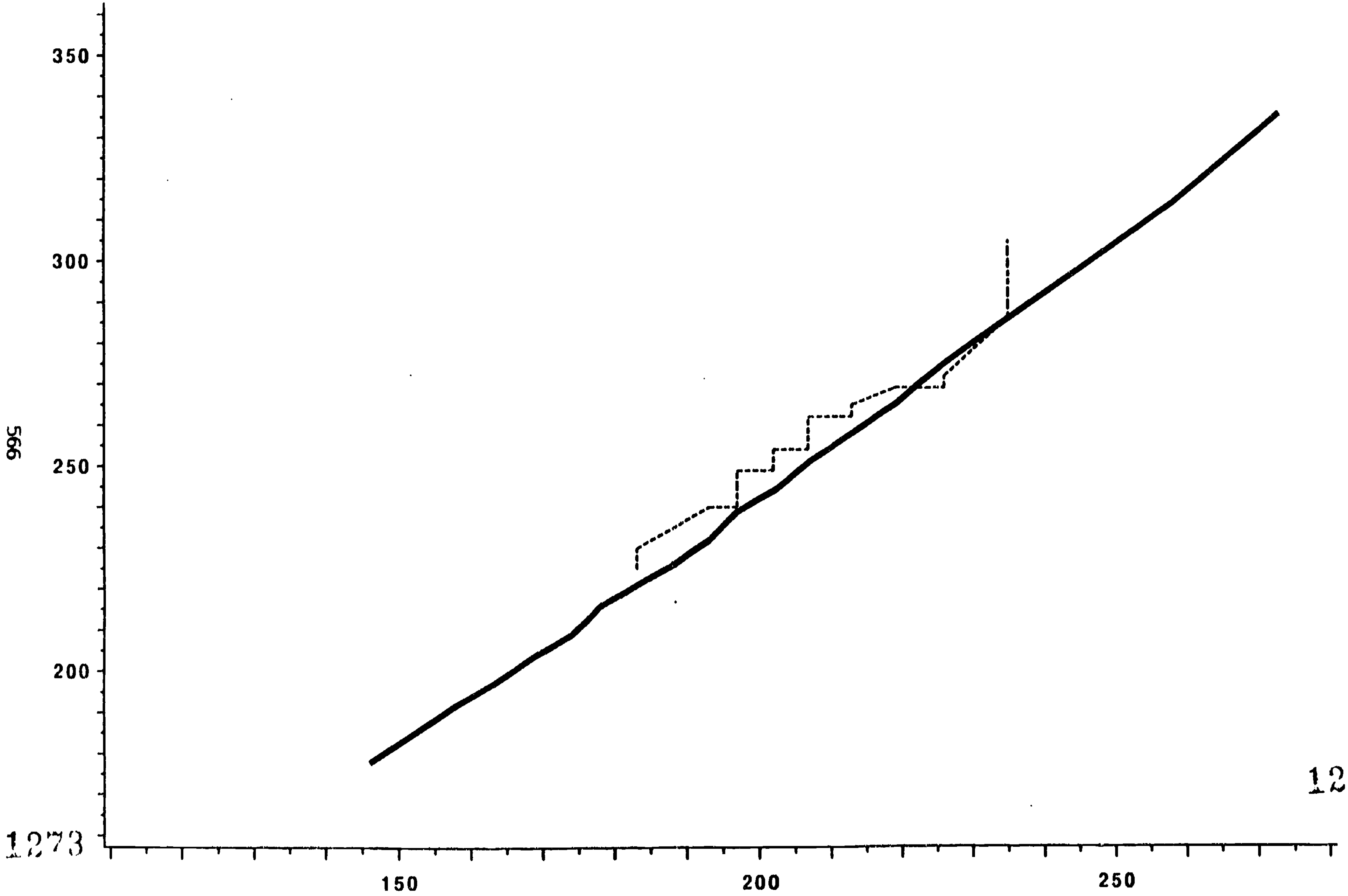
A.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

Students in all three late-exit sites seem to exhibit faster growth rates in the development of skills in mathematics relative to this norming population (site D is barely above), even when assessed in English at spring first grade and in Spanish at spring kindergarten (see Figures 180 to 182). This in part reflects the non-verbal nature of mathematics skills at these levels and the ability of limited-English-proficient students to develop mathematics skills at an even faster rate relative to this norming population when their primary language is used. If these growth rates were sustained, in time each group would catch up with the average performance of this norming population. The critical finding is that each group seems to have been progressing and not losing ground relative to this norming population. Presumably reflecting the greater instruction in English, site G students appear to grow the fastest relative to this norming population.

Figure 180

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



CURVE ▬▬▬ NORM -------- LE-D

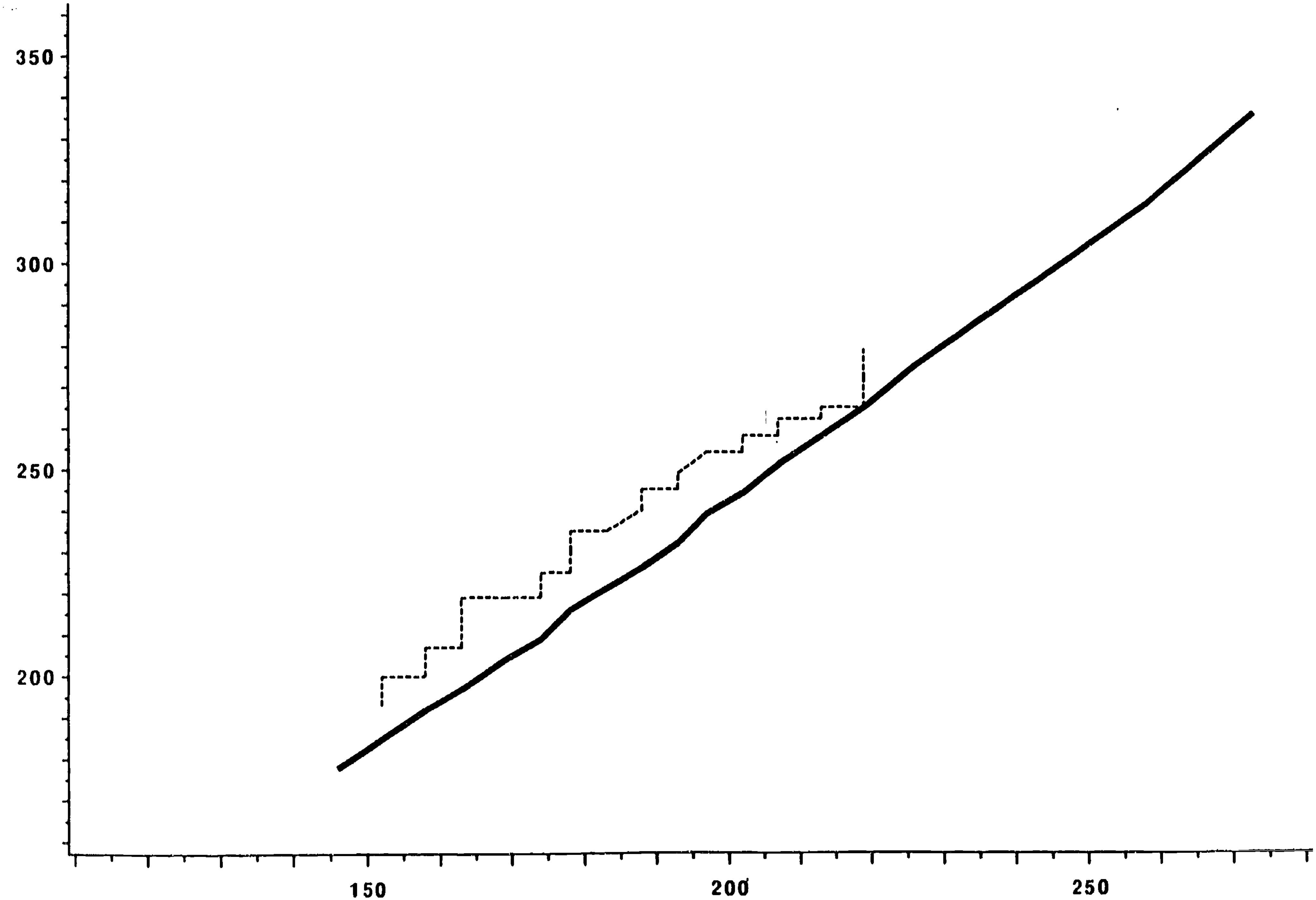
- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, [illegible]l are[illegible]ct to[illegible]ling [illegible]ation.

ERIC

Figure 181

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



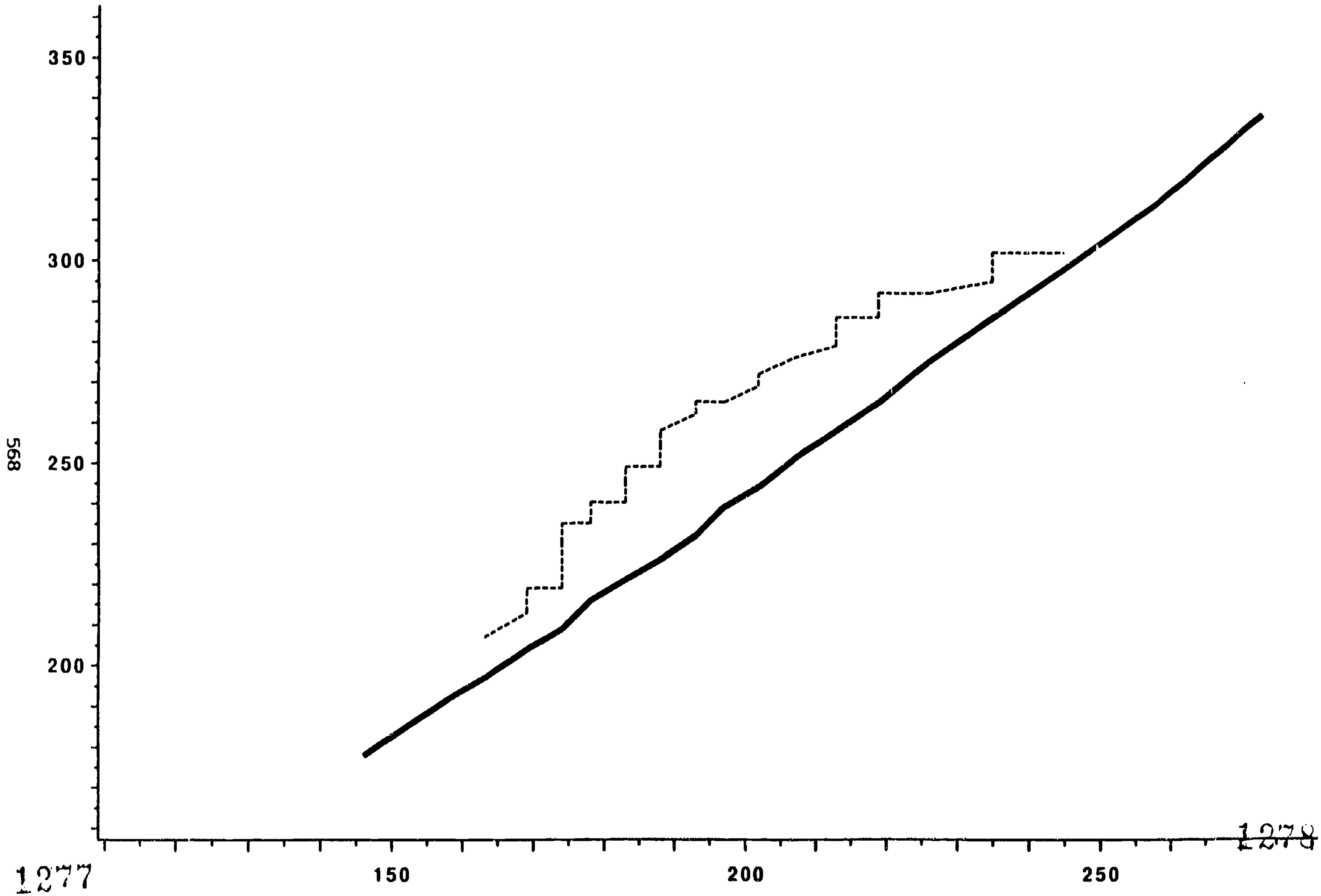
567

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 182

English Math Spring 1 vs. Spanish Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



CURVE ——— NORM - - - - - - - - LE-G

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

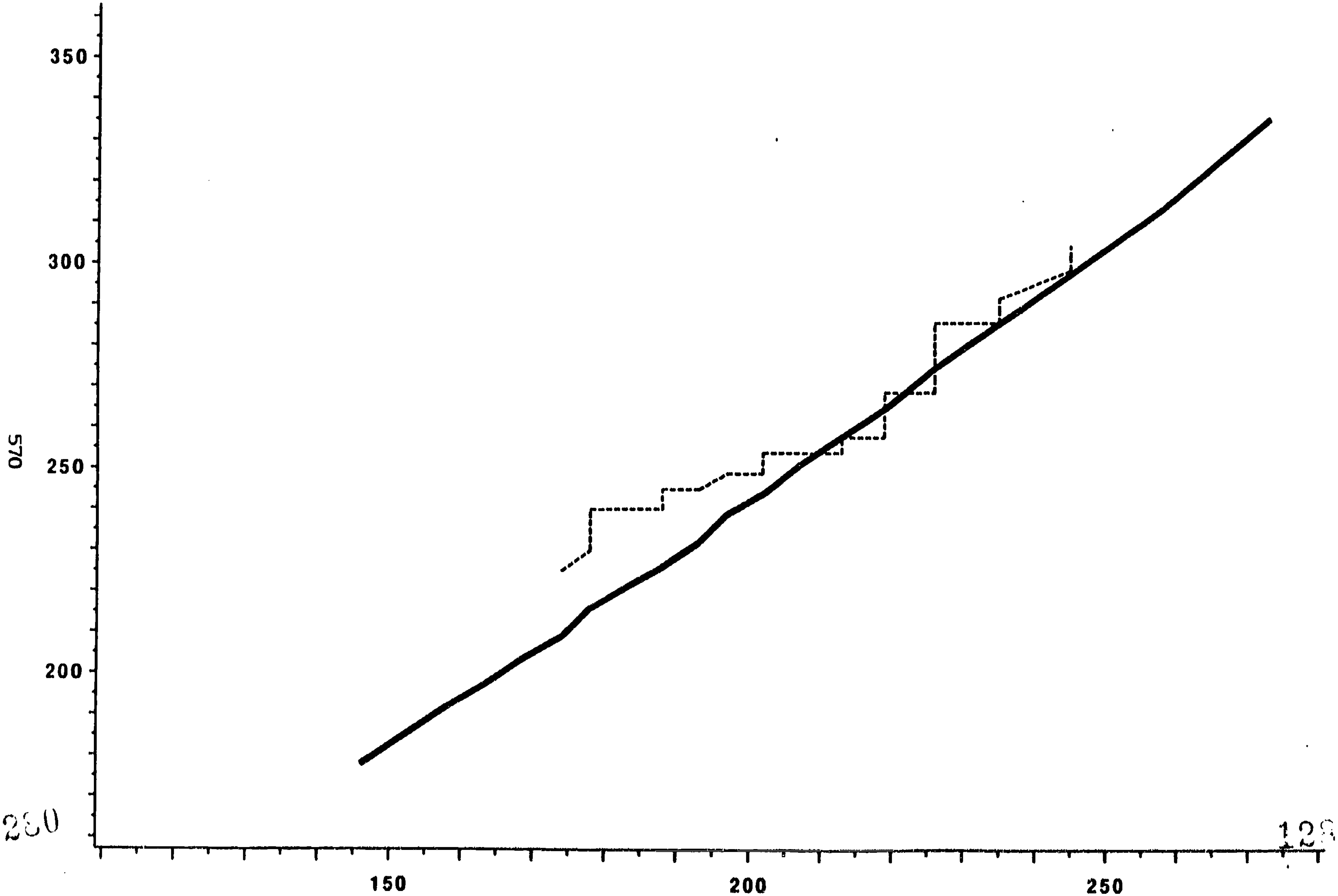
B.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

Figures 183 to 185 suggest that students in all three sites improved their mathematics skills as fast as or at a faster rate than this norming population. As noted earlier, students in site G appeared to show the widest range of skills, with site E students having the lowest and site G the highest spread of initial mathematics skills. Students in sites E and G seemed to grow faster than site D students relative to this norming population. These findings are consistent with the predictions made: when given mathematics instruction in Spanish, students would all show improved skills in this content area, even when assessed in English. What is surprising is their higher growth rate, suggesting that if sustained, the average skills for these students would eventually intersect those of this norming population.

Figure 183

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



1280 1281

BEST COPY AVAILABLE

CURVE ——— NORM - - - - - - - - L.E-D

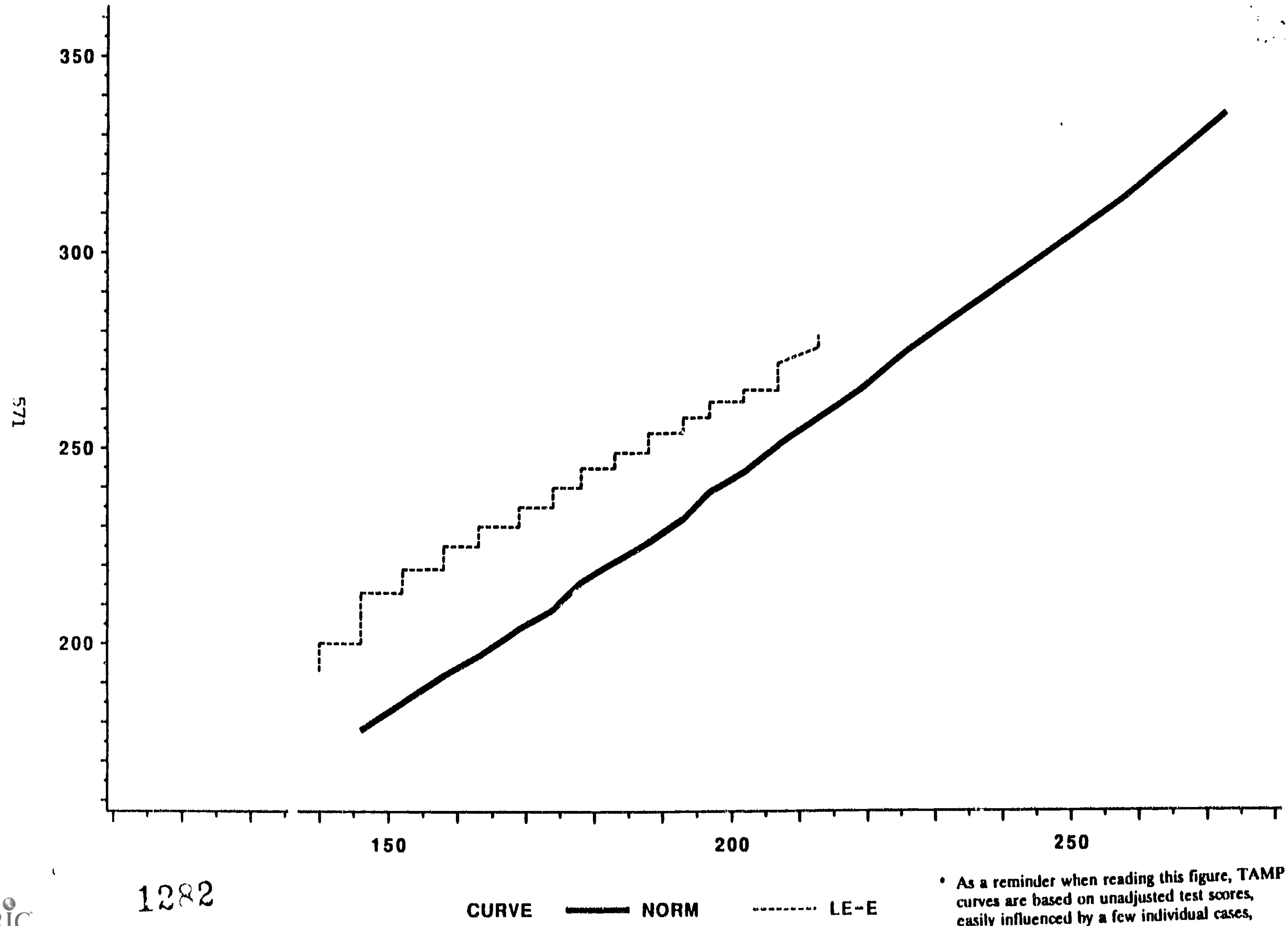
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

Figure 184

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E

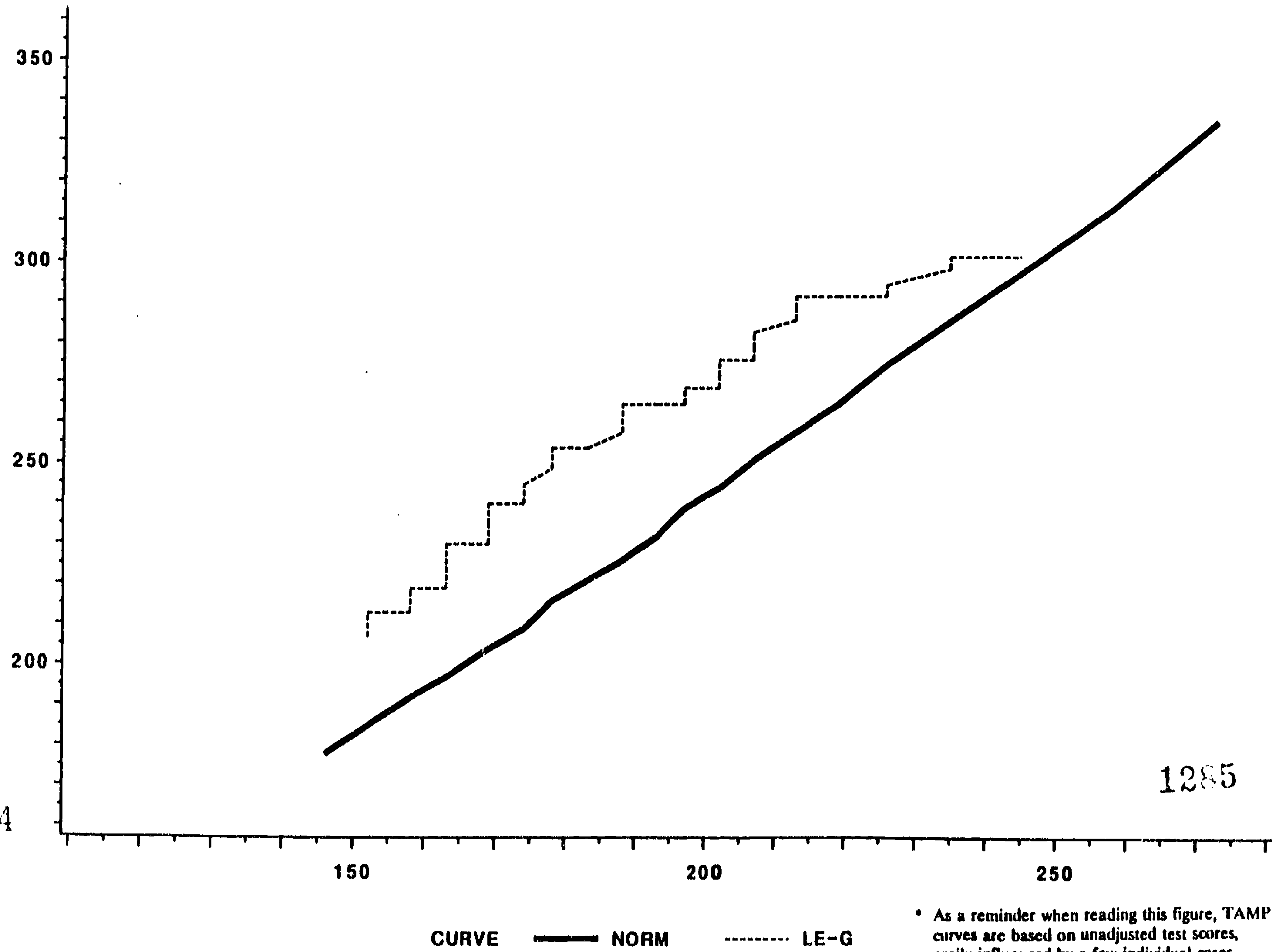


- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 185

English Math Spring 1 vs. English Math Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, [illegible]

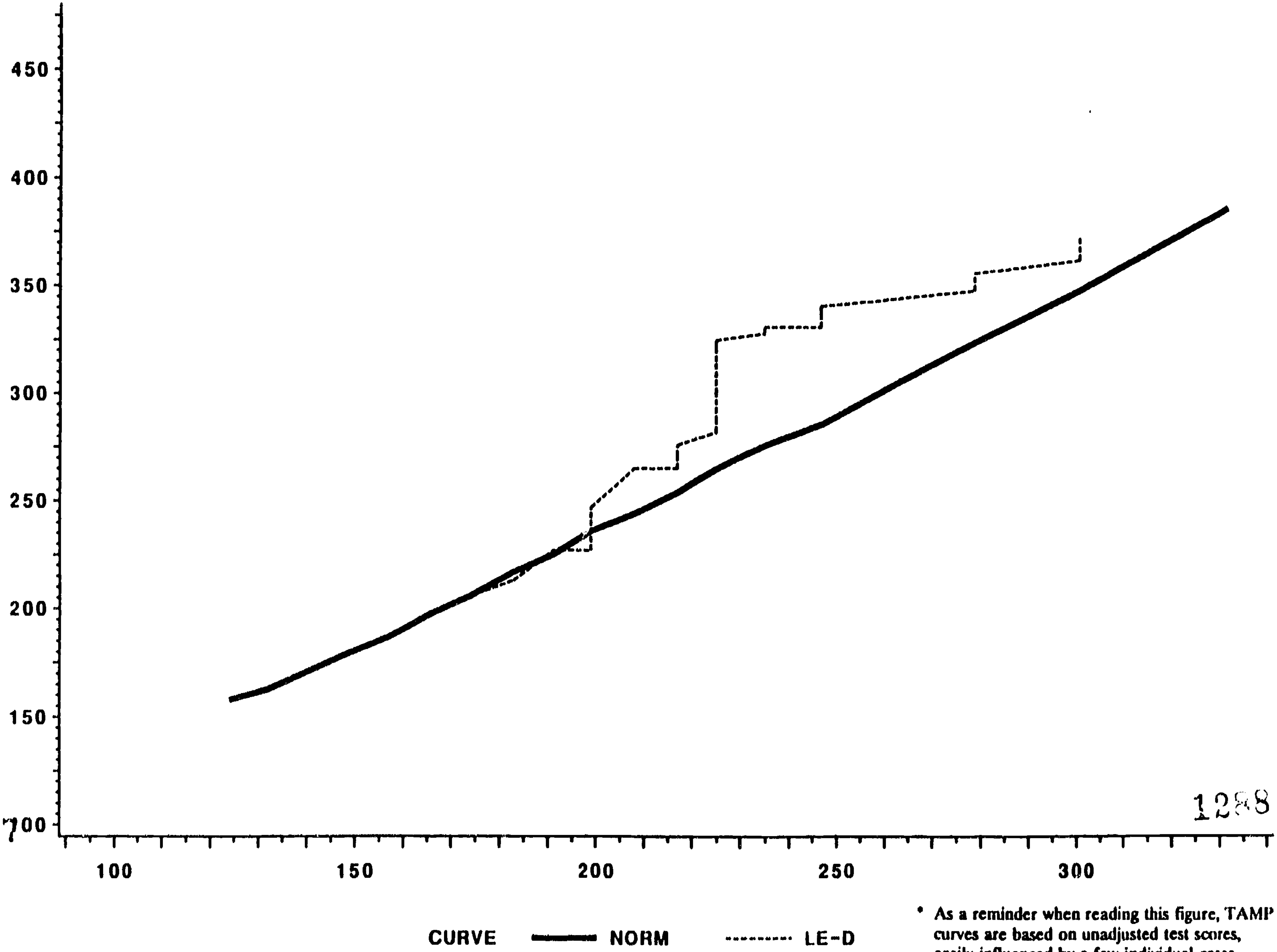
C. **Grade Span:**	**Kindergarten**	**to**	**First Grade**
Test Date:	**Spring**	**to**	**Spring**
Language:	**Spanish**	**to**	**English**
Content:	**Language**	**to**	**Language**

Students in site D appear to have learned English language skills at a faster rate relative to this norming population and to site E and site G students (see Figures 186 to 188). Both site E and site G students seemed to learn their English language skills at about the same rate as this norming population. These curves also show that students at site E ended kindergarten and first grade with lower English language skills than students in site D or G. Students in site G had the broadest range of skills as compared to site D and site E students. Noteworthy is the finding that in spite of the more limited amount of instruction in English, contrary to expectations, students in all three sites seemed to acquire English language skills as fast as or slightly faster than this norming population, given where they stood at spring kindergarten in Spanish.

Figure 186

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



574

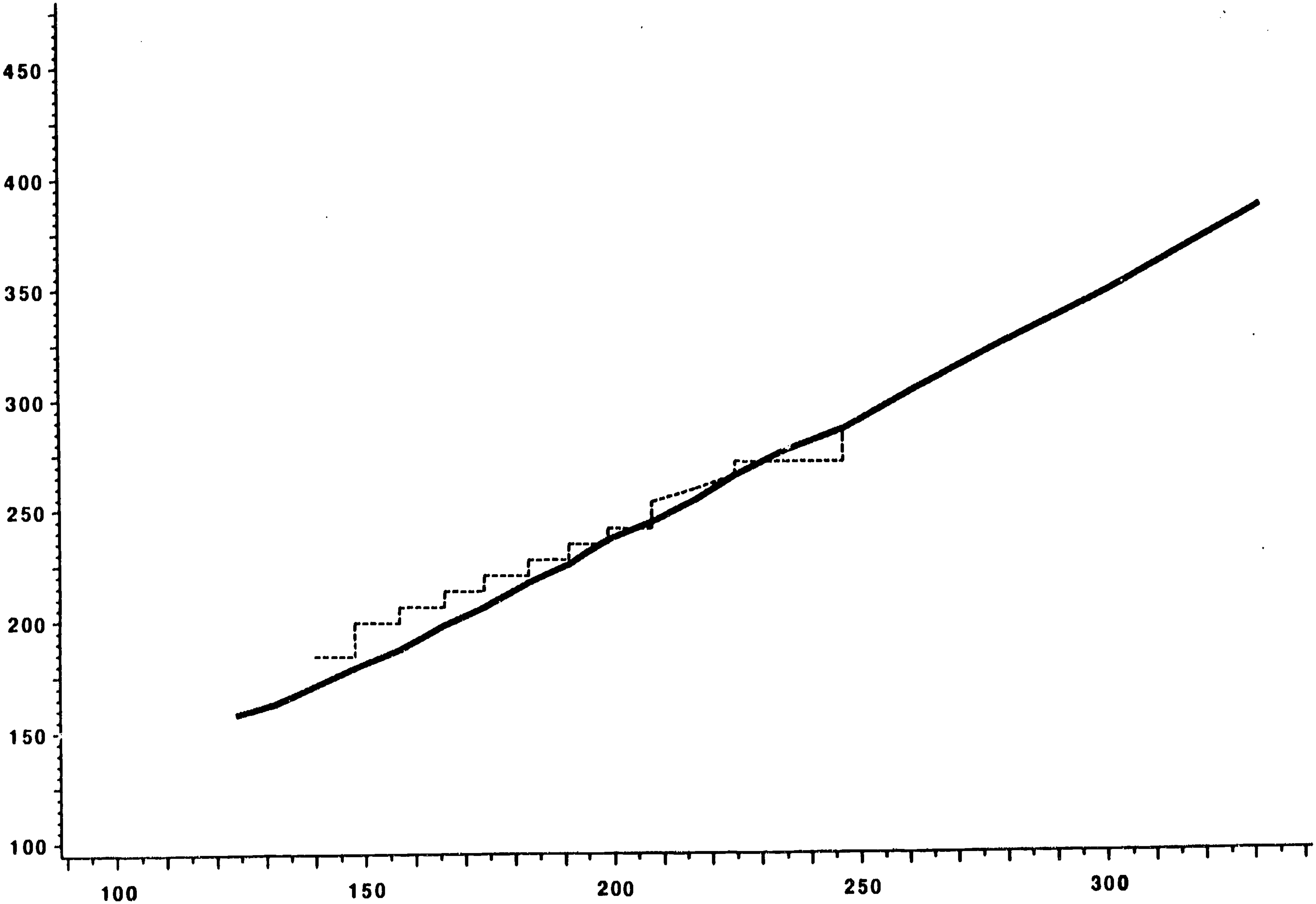
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

ERIC

Figure 187

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



575

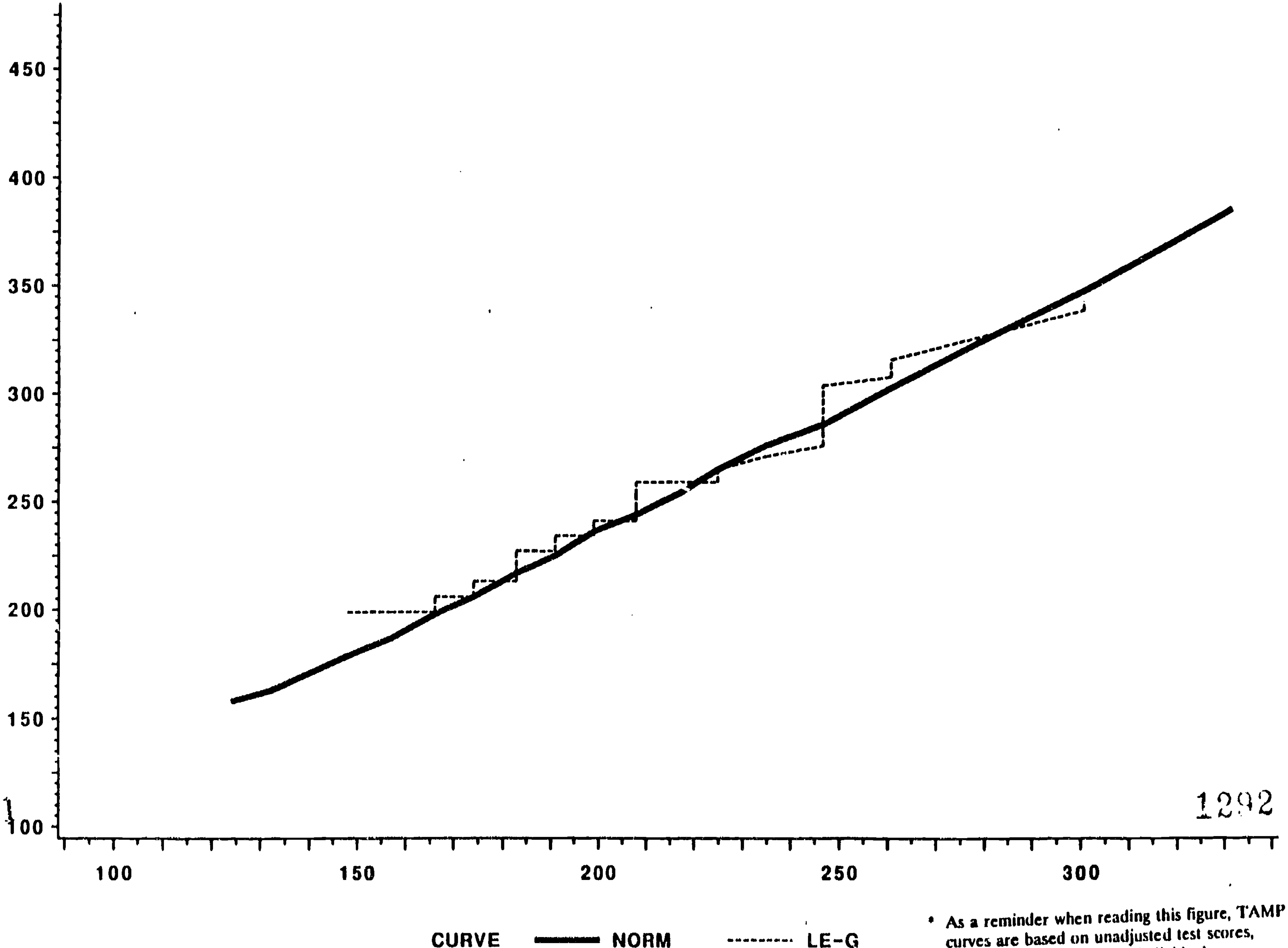
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 188

English Language Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



576

1291 1292

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, ...

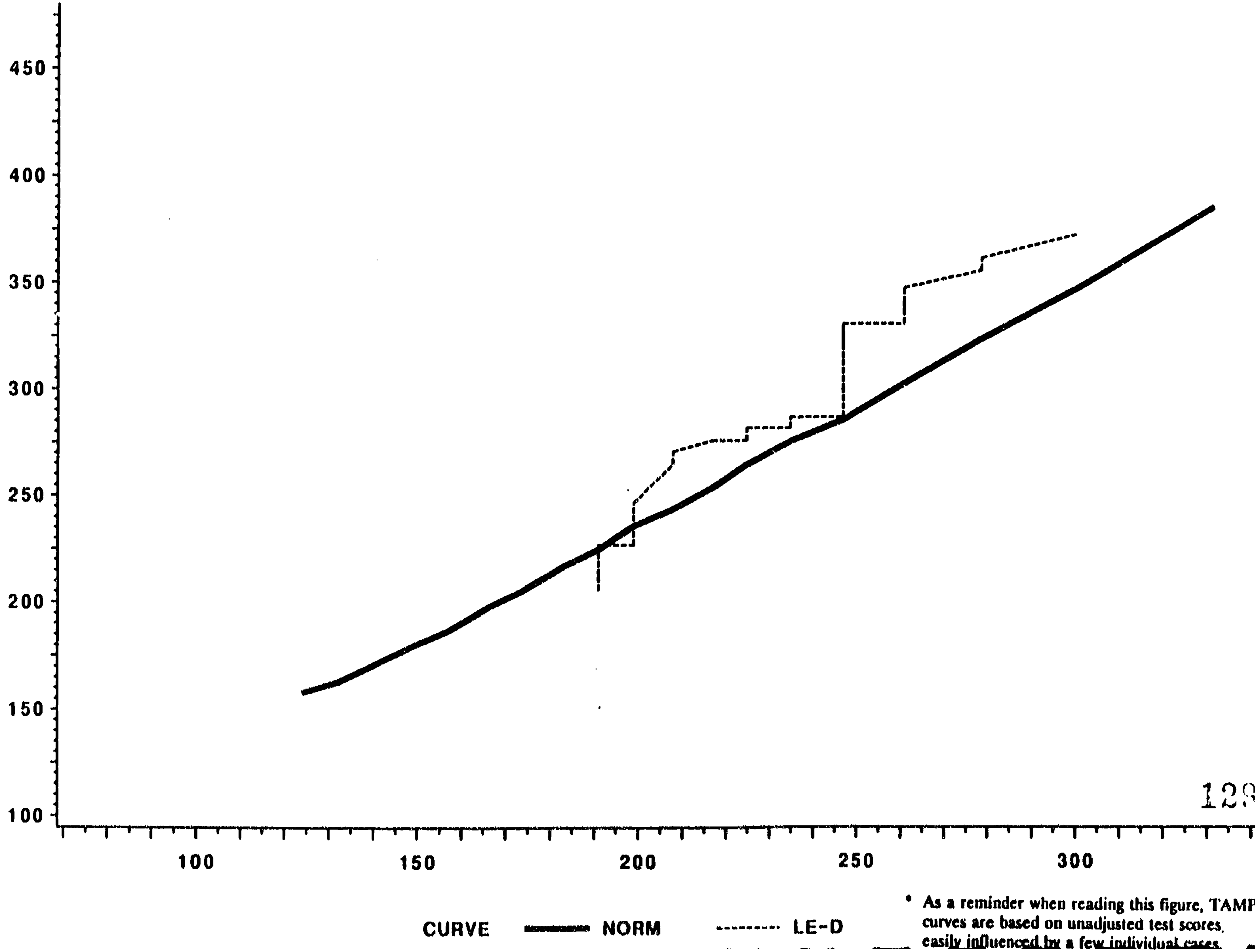
D.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Language**	**to**	**Language**

Figures 189 to 191 suggest that students in all three late-exit sites showed an increase in their English language skills that was as high as or higher than this norming population, given where they stood at spring kindergarten. Site E students seemed to show the lowest distribution of spring first grade scores, followed by site G students having the widest range of spring first grade scores. Site D students seemed to have a narrower but higher range of spring first grade scores. These findings are surprising in that they were not predicted by the late-exit model. The late-exit model posited that growth would be modest at best for each site given the limited instruction in English.

Figure 189

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

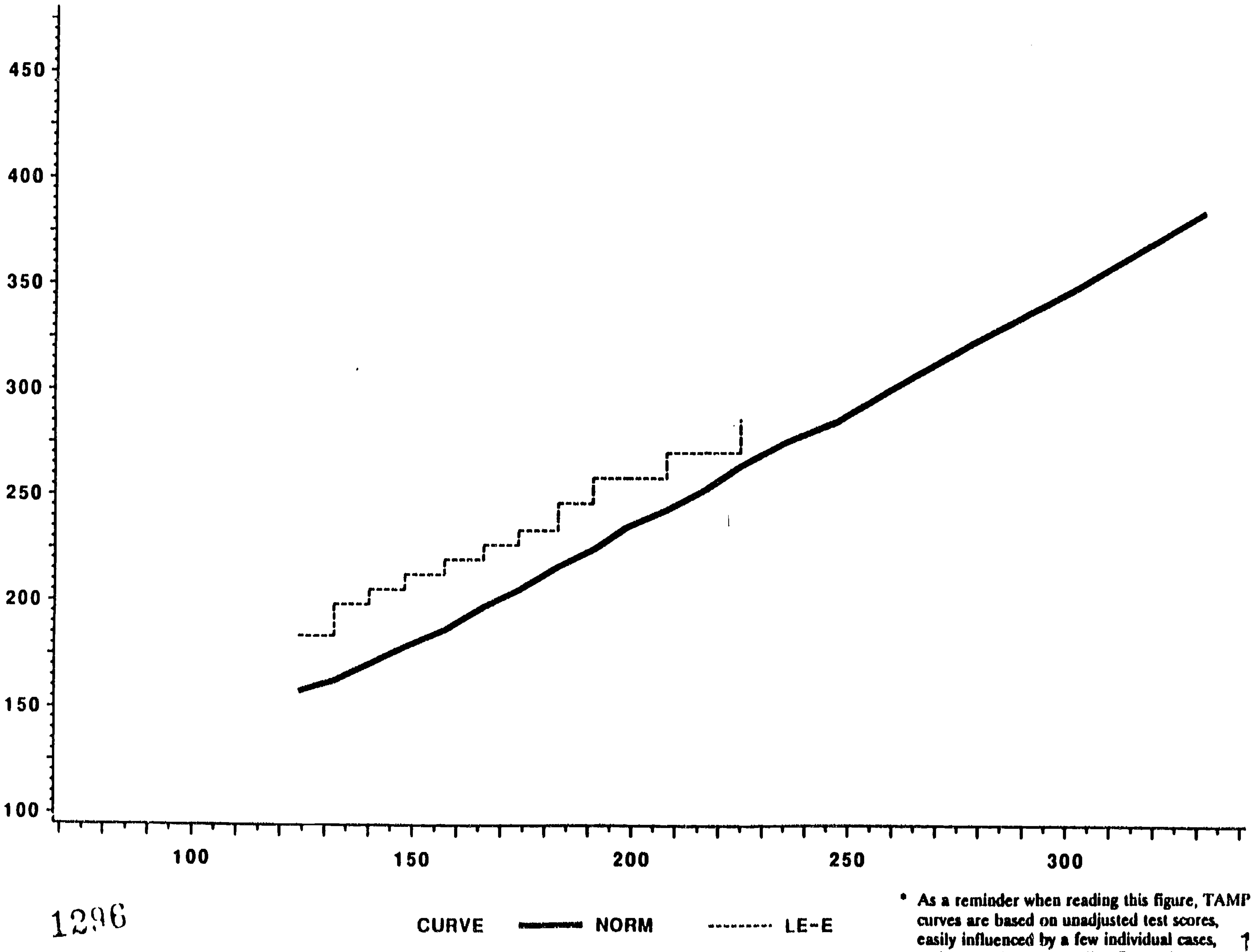
578

1284 1285

Figure 190

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



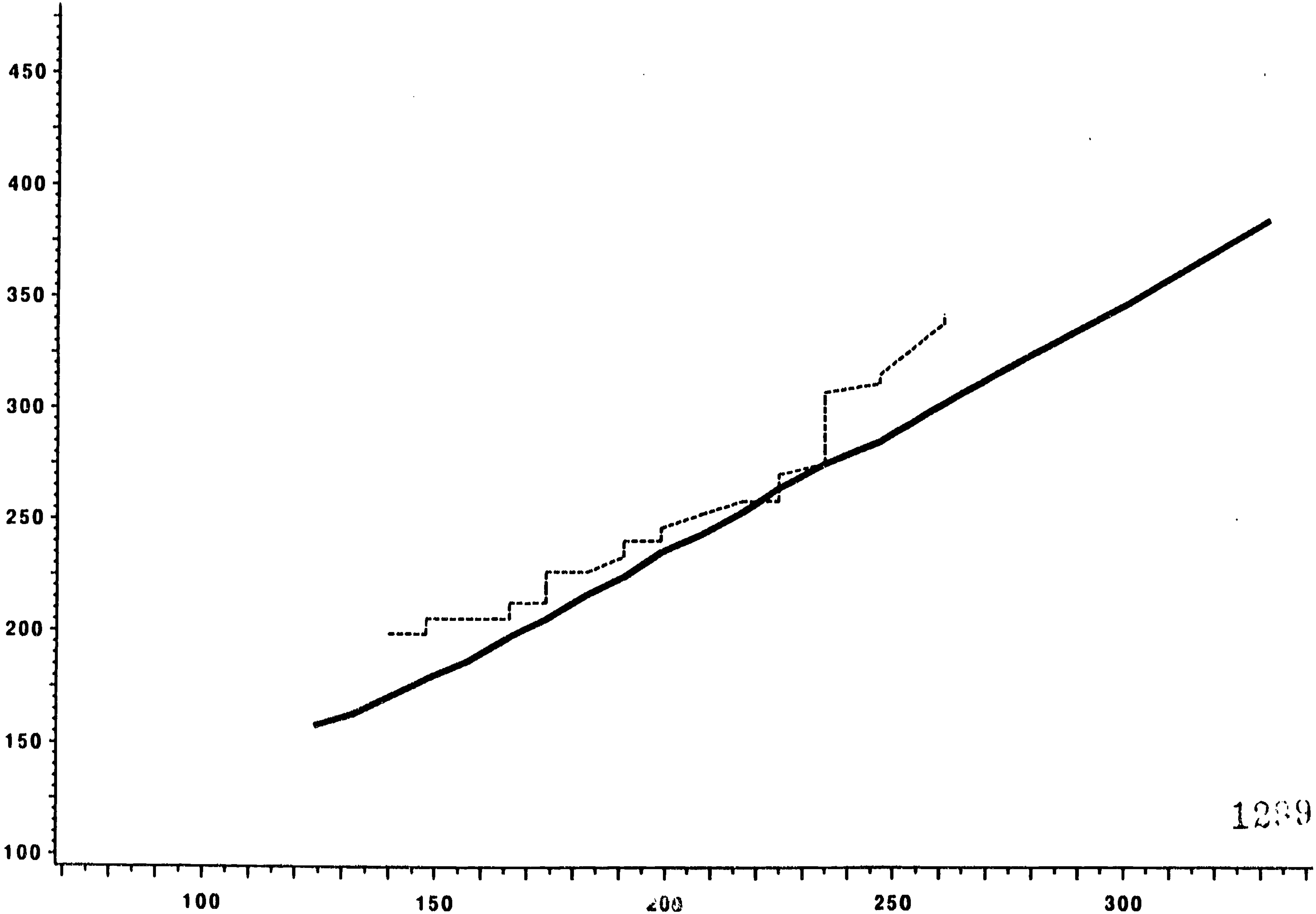
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 191 *

English Language Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



CURVE ——— NORM -------- LE-G

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

1288

1289

ERIC

E. Grade Span:	**Kindergarten**	**to**	**First Grade**
Test Date:	**Spring**	**to**	**Spring**
Language:	**Spanish**	**to**	**English**
Content:	**Language**	**to**	**Reading**

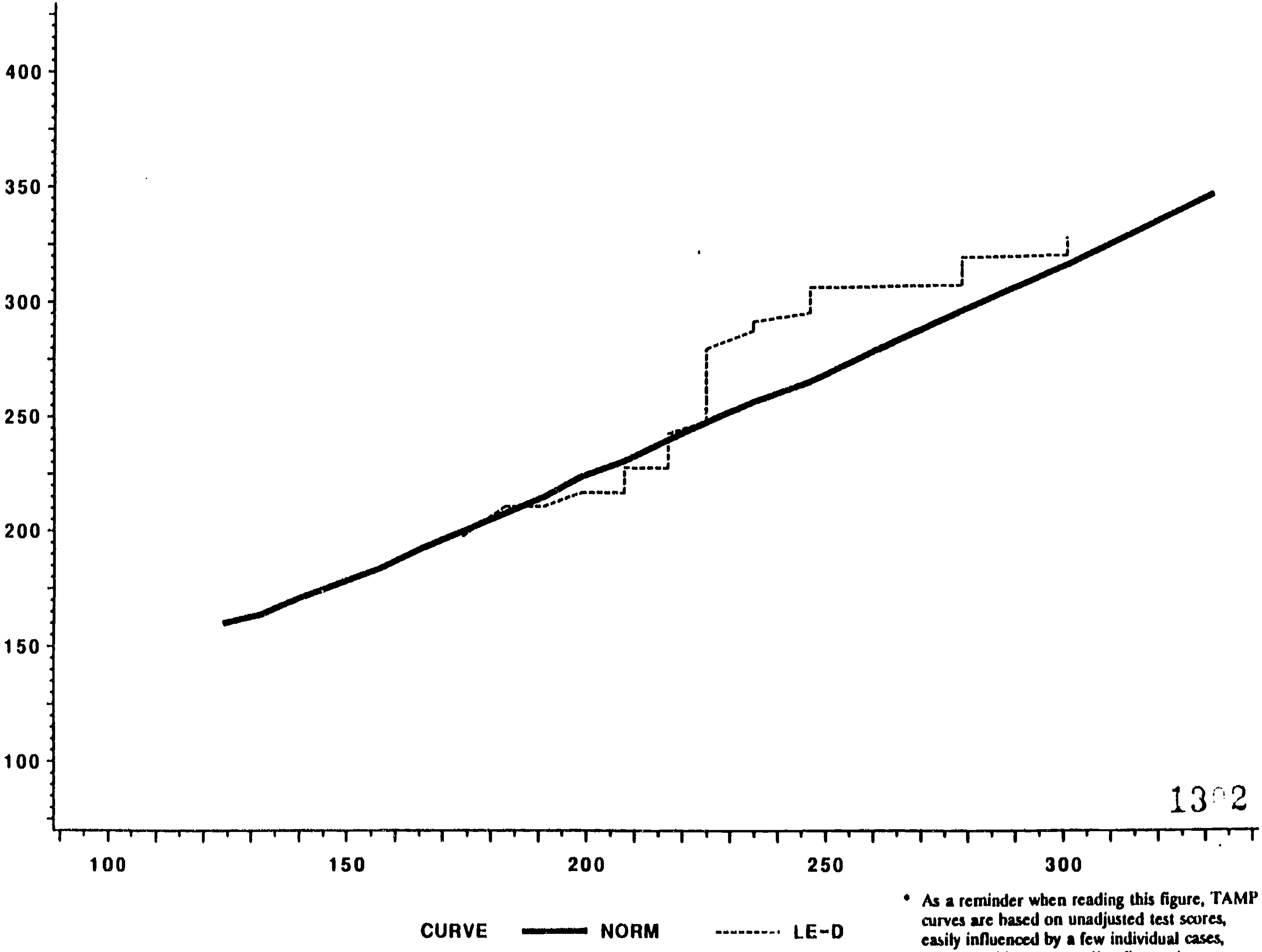
Figures 192 to 194 suggest that students in sites D and G ended their kindergarten year with roughly the same range of Spanish language skills relative to this norming population. Site E students appeared to end their kindergarten year with lower Spanish language skills, reflecting their more modest home and community backgrounds. These figures suggest that site E students increased their English reading skills at roughly the same rate relative to this norming population, whereas site D students appear to have learned to read in English at the same or slightly faster rate, and site G students at a faster rate. Students in all three sites performed better than would have been predicted from the late-exit model given the smaller amount of instruction provided in English reading. This tends to support the notion that improving or strengthening the primary language skills facilitates the acquisition of second language skills.

ERIC

Figure 192

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



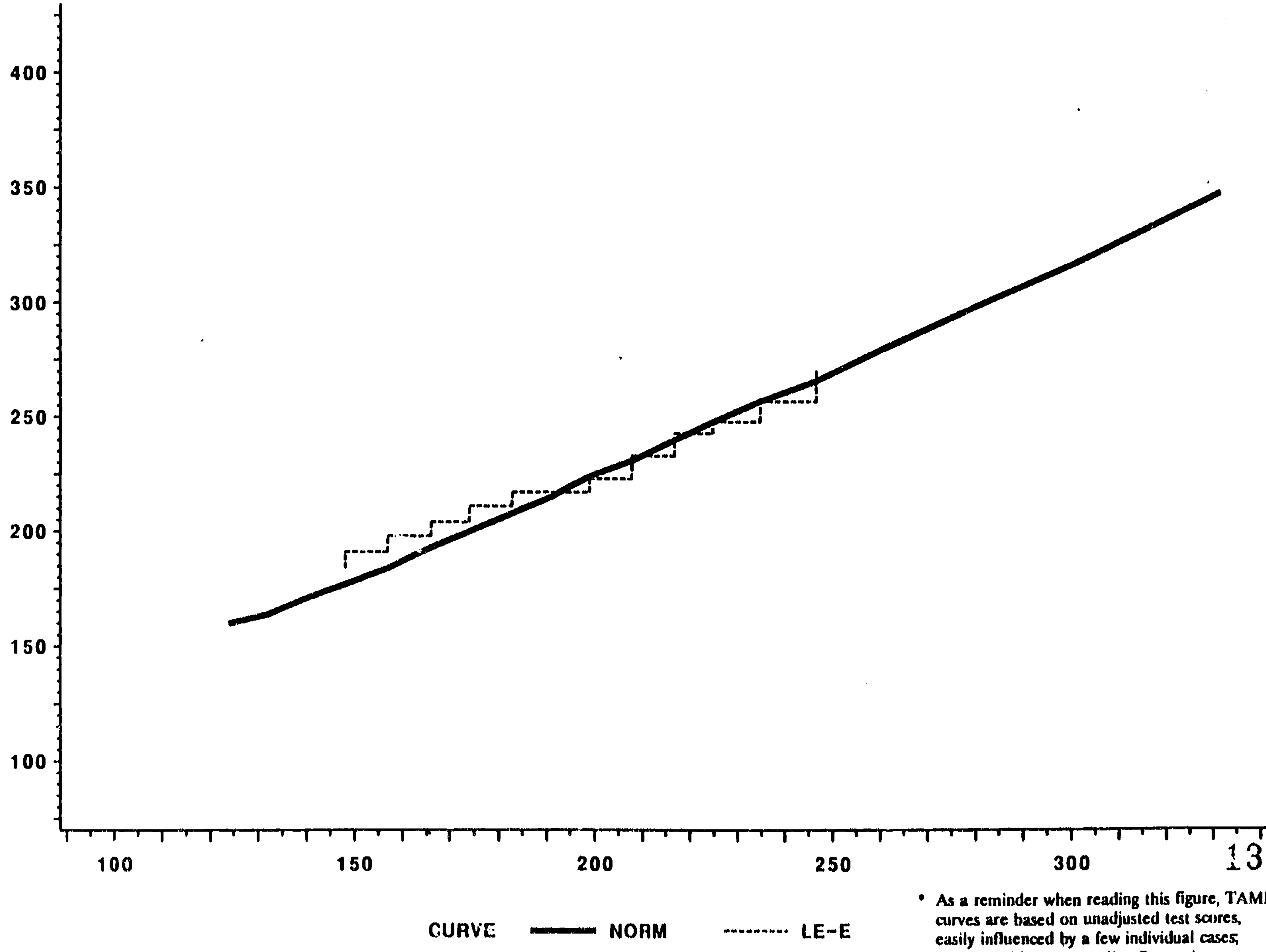
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuations.

ERIC

Figure 193

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



583

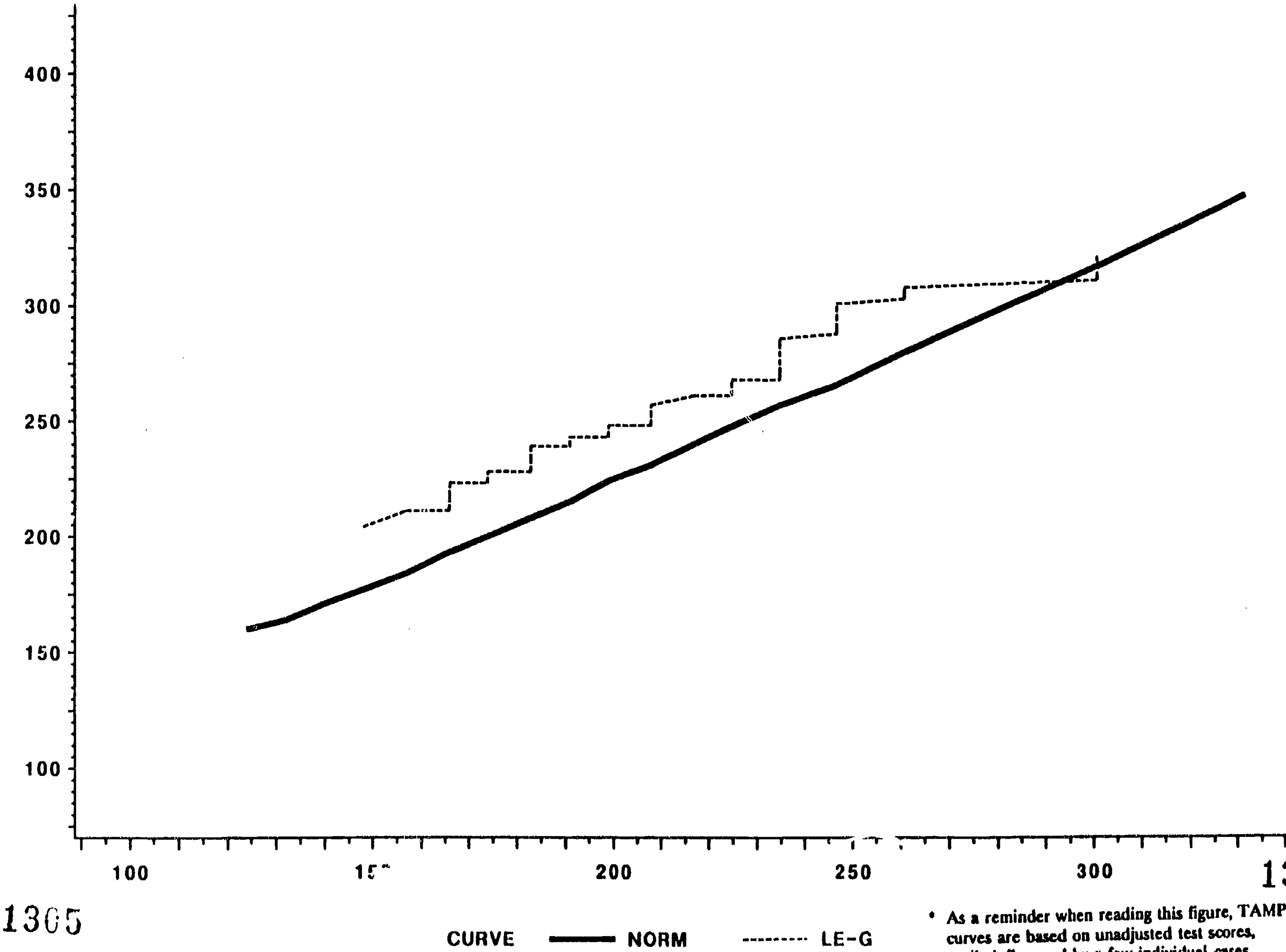
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases; and are subject to sampling fluctuation.

ERIC

Figure 194

English Reading Spring 1 vs. Spanish Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



ERIC

• As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

F.

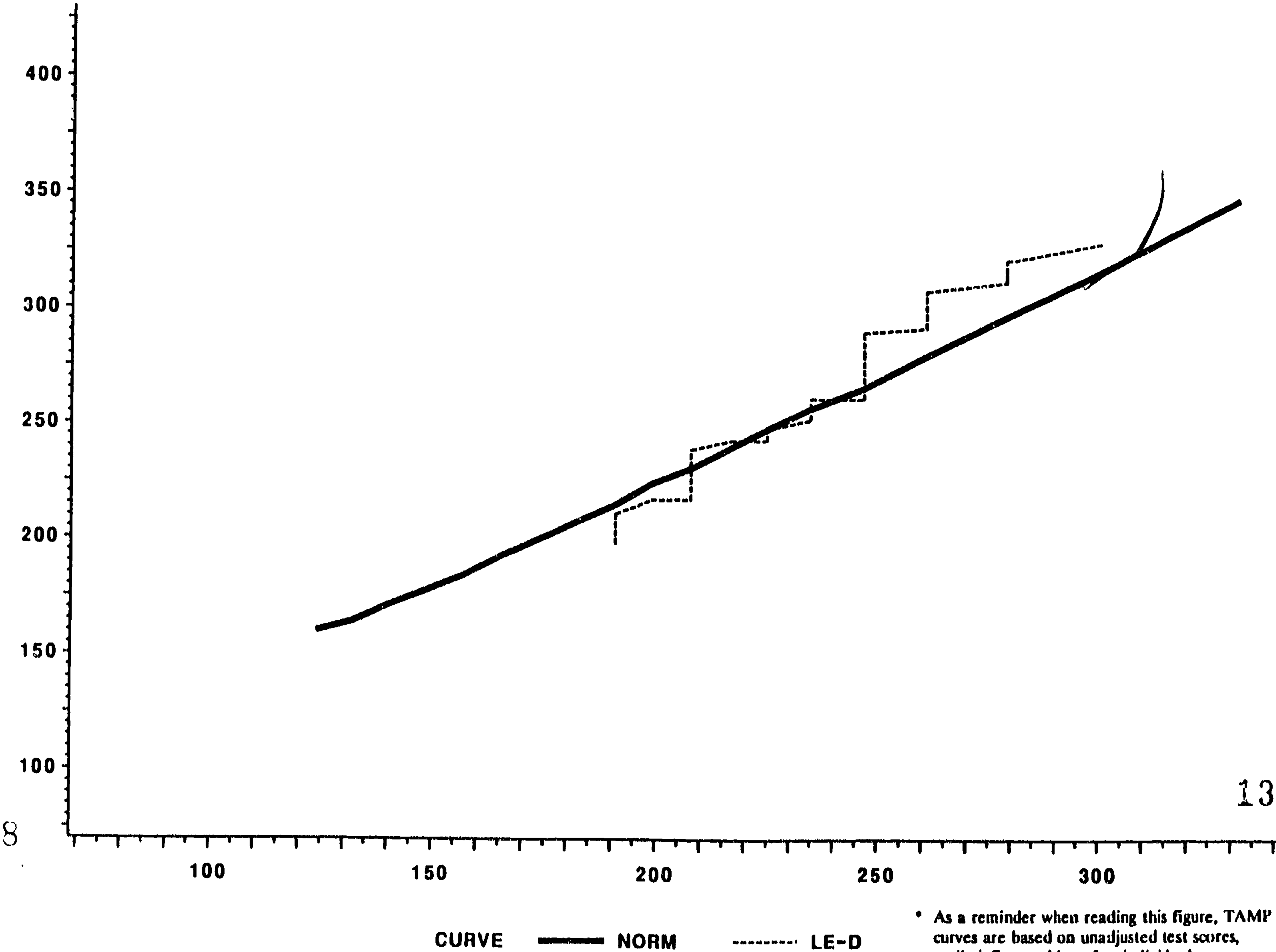
Grade Span:	Kindergarten	to	First Grade
Test Cycle:	Spring	to	Spring
Language:	English	to	English
Content:	Language	to	Reading

As we have observed before, students in site D appeared to end their kindergarten year with the highest range of English language skills relative to this norming population, followed by site G and site E students, respectively (see Figures 195 to 197). Site E and site G students seemed to learn to read English at a faster rate relative to this norming population, whereas site D students appeared to learn to read in English almost at the same rate as this norming population. Once again, these growth rates are better than would have been predicted from the late-exit instructional model. The original predictions suggested that the growth in language skills would be minimal given the limited amount of English instruction. These patterns are also supportive of the notion of providing special support services to language-minority students; if these growth rates were sustained over time, the distribution of achievement for these student would approximate this norming population.

Figure 195

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



586

13‵8

1309

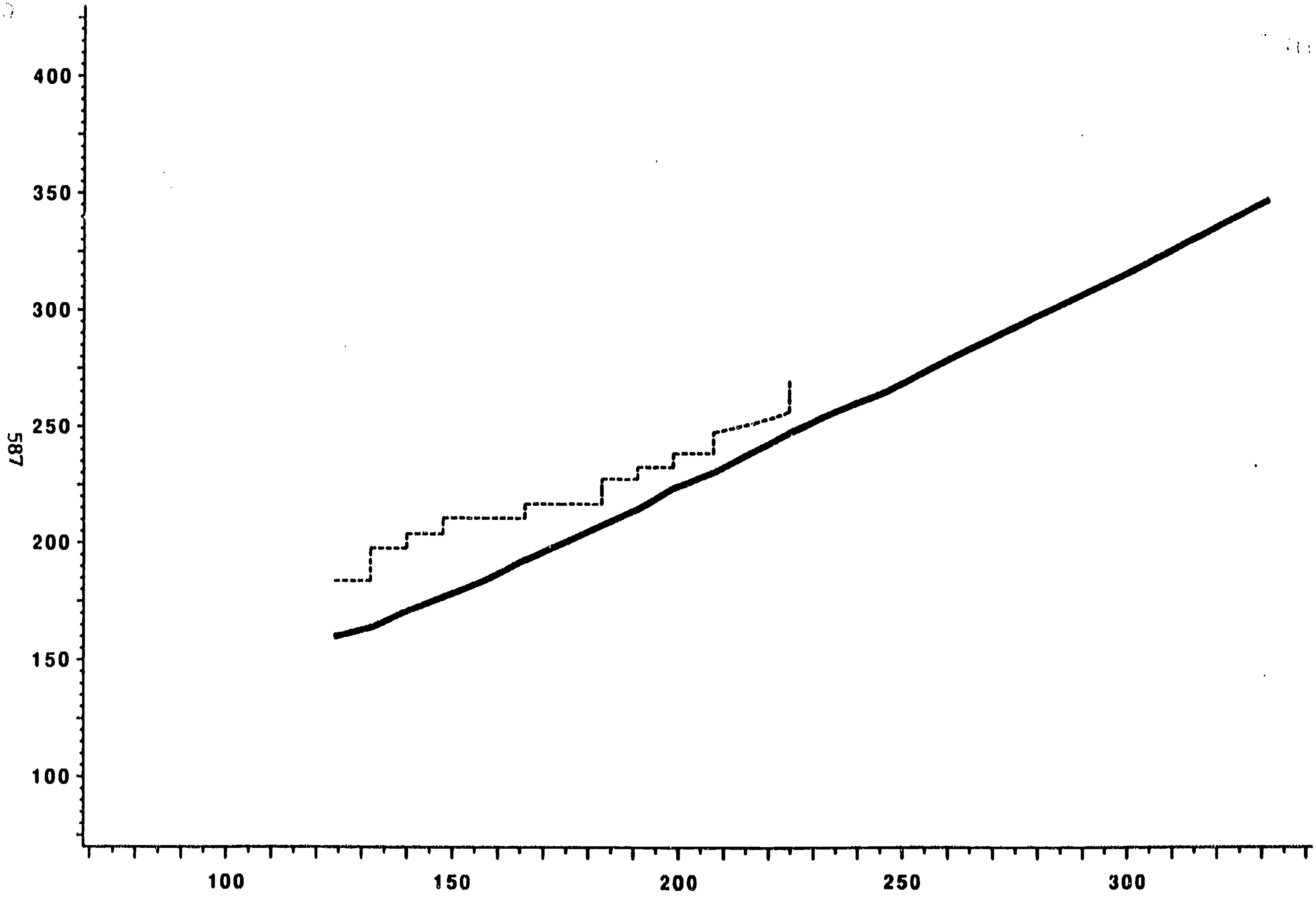
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

ERIC

Figure 196

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



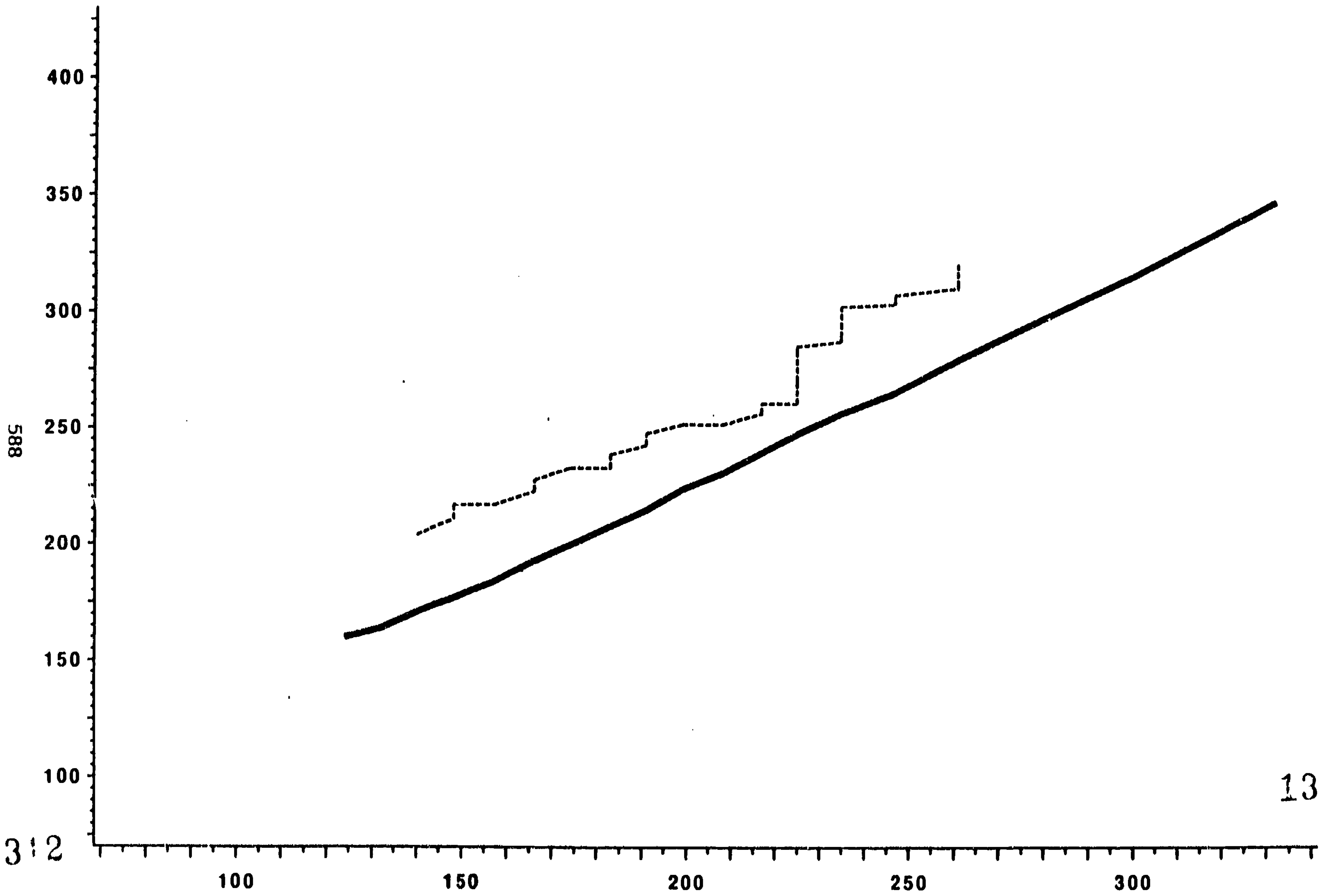
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 197

English Reading Spring 1 vs. English Language Spring K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



588

13!2

13!3

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

3. Fall Kindergarten to Spring First Grade:

A.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

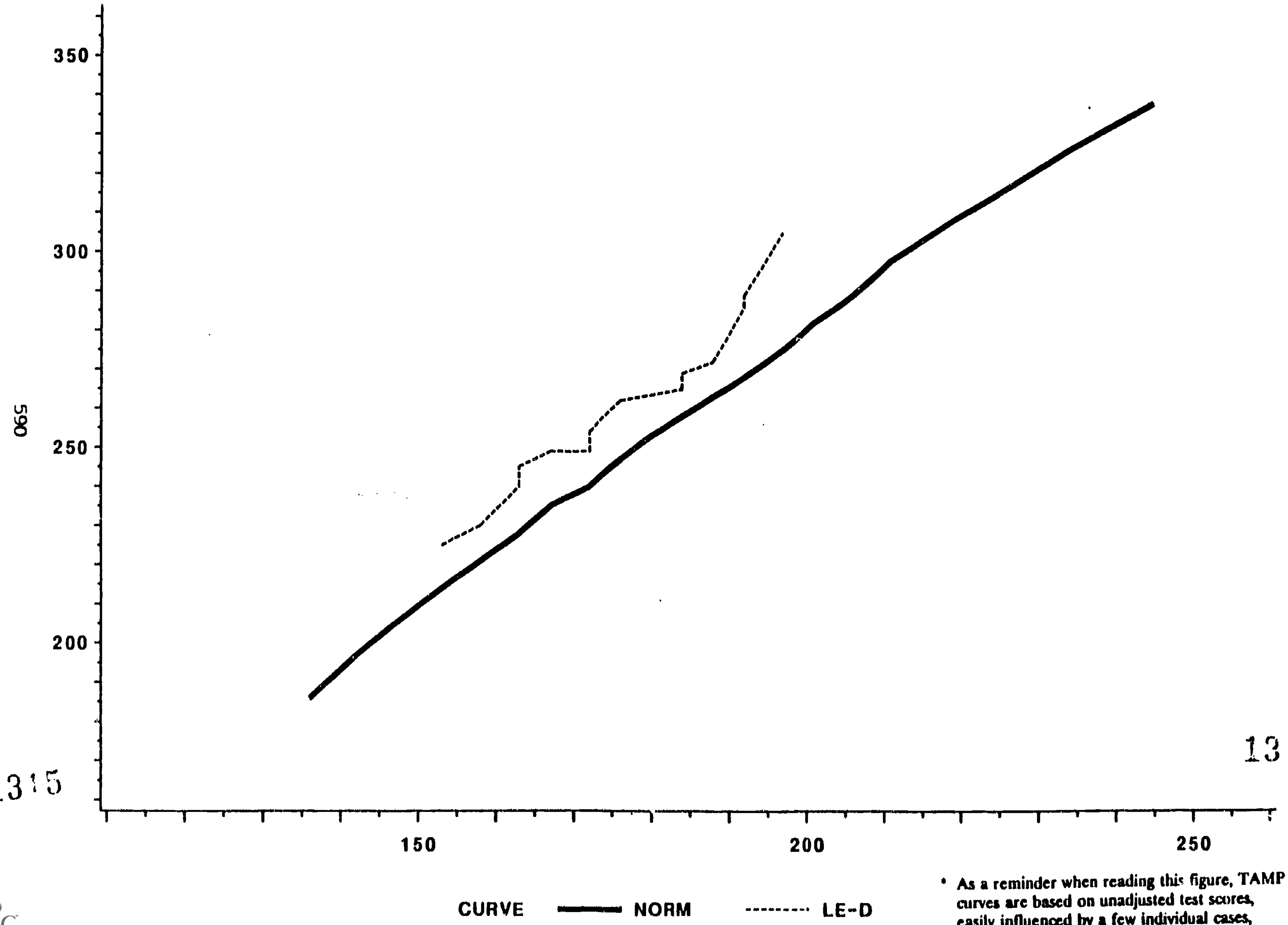
As noted above, site D students began kindergarten with slightly higher skills in mathematics than site E students relative to this norming population (Figures 198 and 199). However, both groups of students seemed to learn at a faster rate relative to this norming population. The growth in mathematics skills appears to be comparable between the two groups. These results are better than expected given the predictions made from the late-exit model which projected growth in the content areas to be slower if tested in English.

ERIC

Figure 198

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



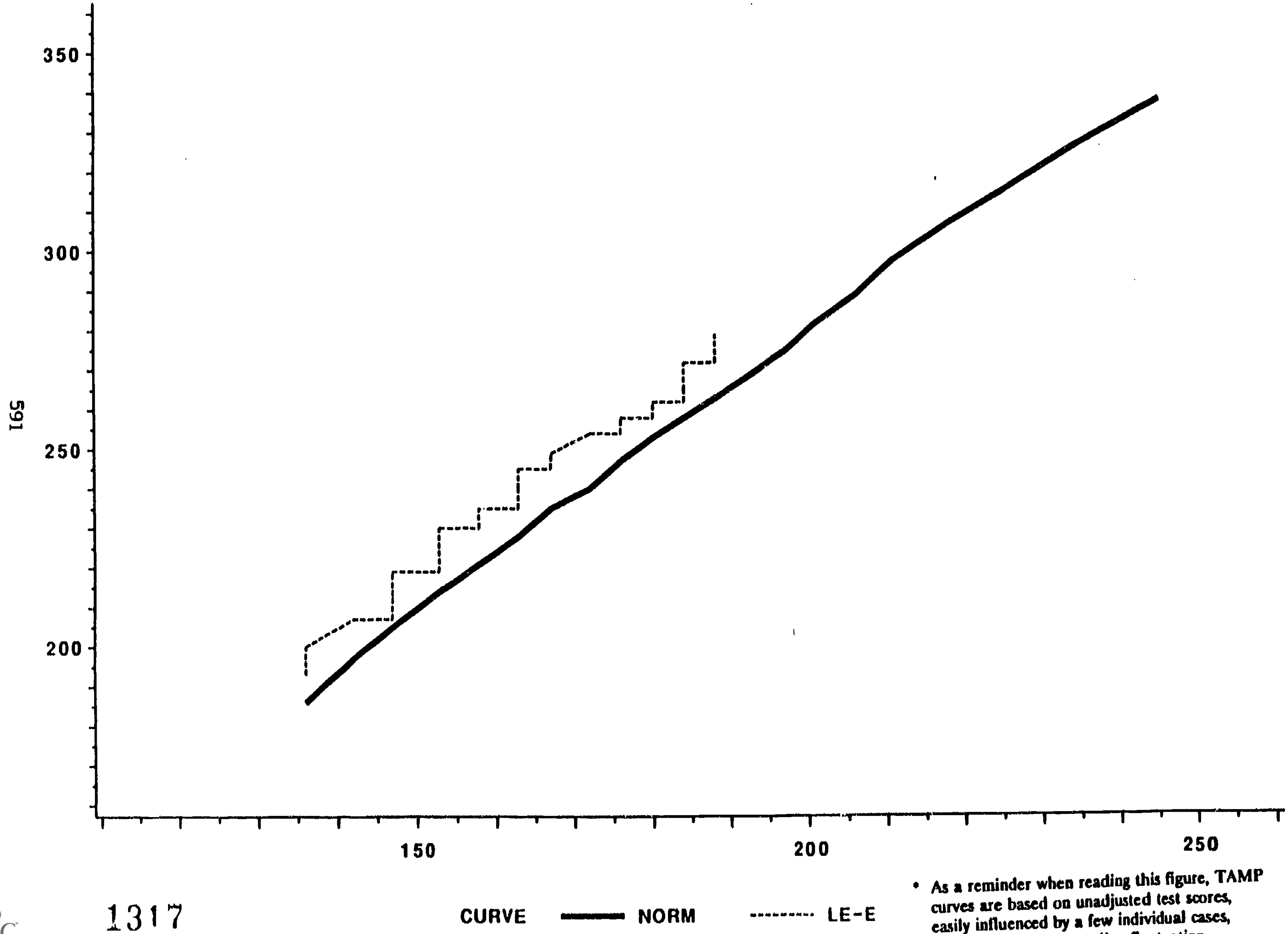
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

ERIC

Figure 199

English Math Spring 1 vs. Spanish Math Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

B. Grade Span:	**Kindergarten**	**to**	**First Grade**
Test Date:	**Fall**	**to**	**Spring**
Language:	**Spanish**	**to**	**English**
Content:	**Language**	**to**	**Language**

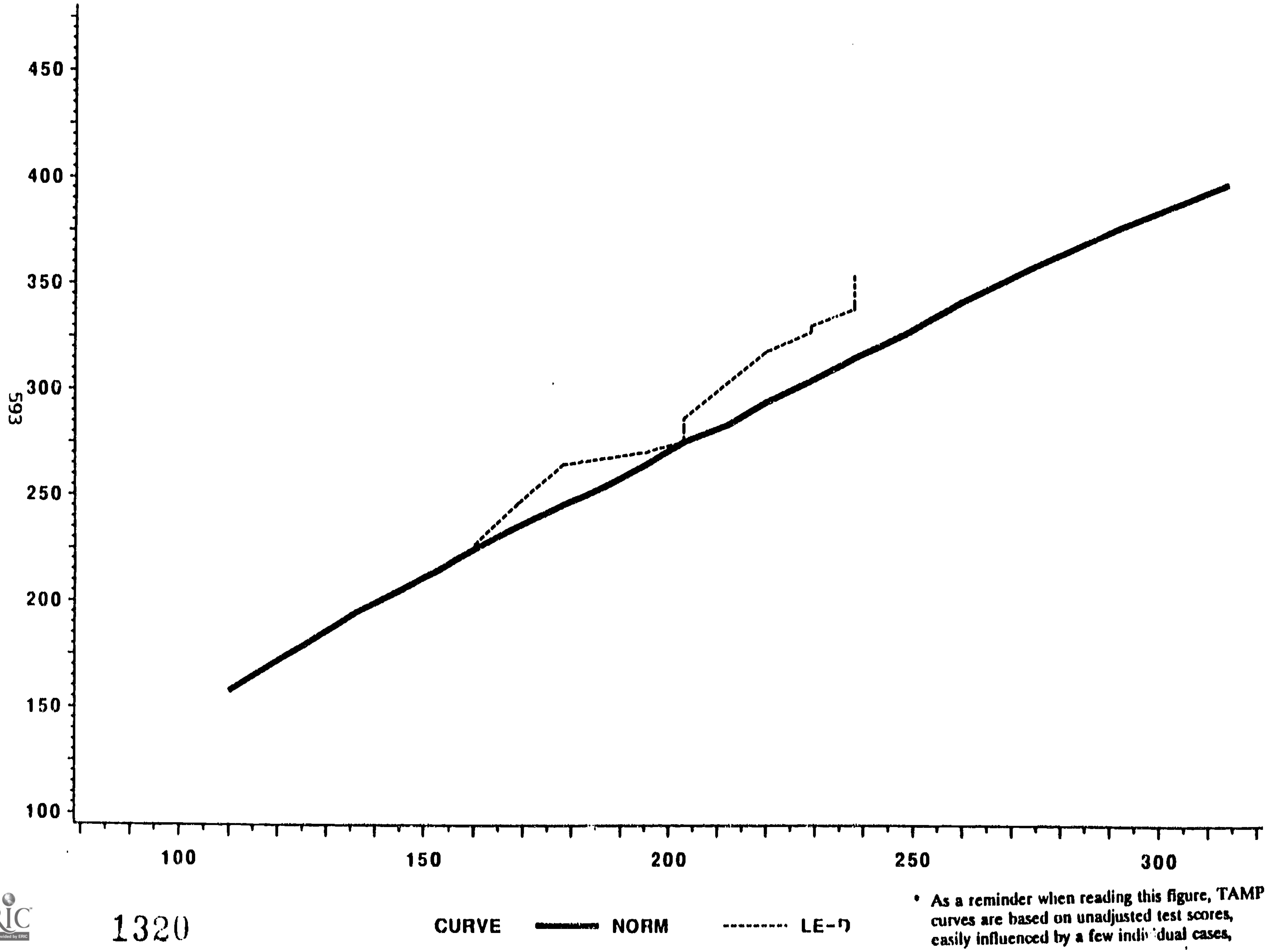
Site D students appear to have acquired English language skills at a slightly faster rate than site E students relative to this norming population (see Figures 200 and 201). Site E students entering with lower Spanish language skills appear to have learned English language skills at the same rate as this norming population, whereas site E students with higher entry-level Spanish language skills seem to have developed their English language skills at a slower rate than this norming population. This last observation is consistent with the late-exit model in that late-exit students were predicted to develop these language skills at a slower rate given their limited instruction in English at this grade level. At first glance, the higher than expected performance by site D students might seem inconsistent with this late-exit hypothesis. However, as seen in Figures 176 and 177, both site D and site E students ended their kindergarten year with roughly the same range of Spanish language skills. This better than expected performance by site D students in first grade might suggest greater support for primary language. This explanation might have resulted in higher primary language skills among site D students which, according to the late-exit model, might have facilitated the learning of second language skills. These results could also reflect the higher socioeconomic/educational level of the families and community in site D compared to site E.

ERIC

Figure 200

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D

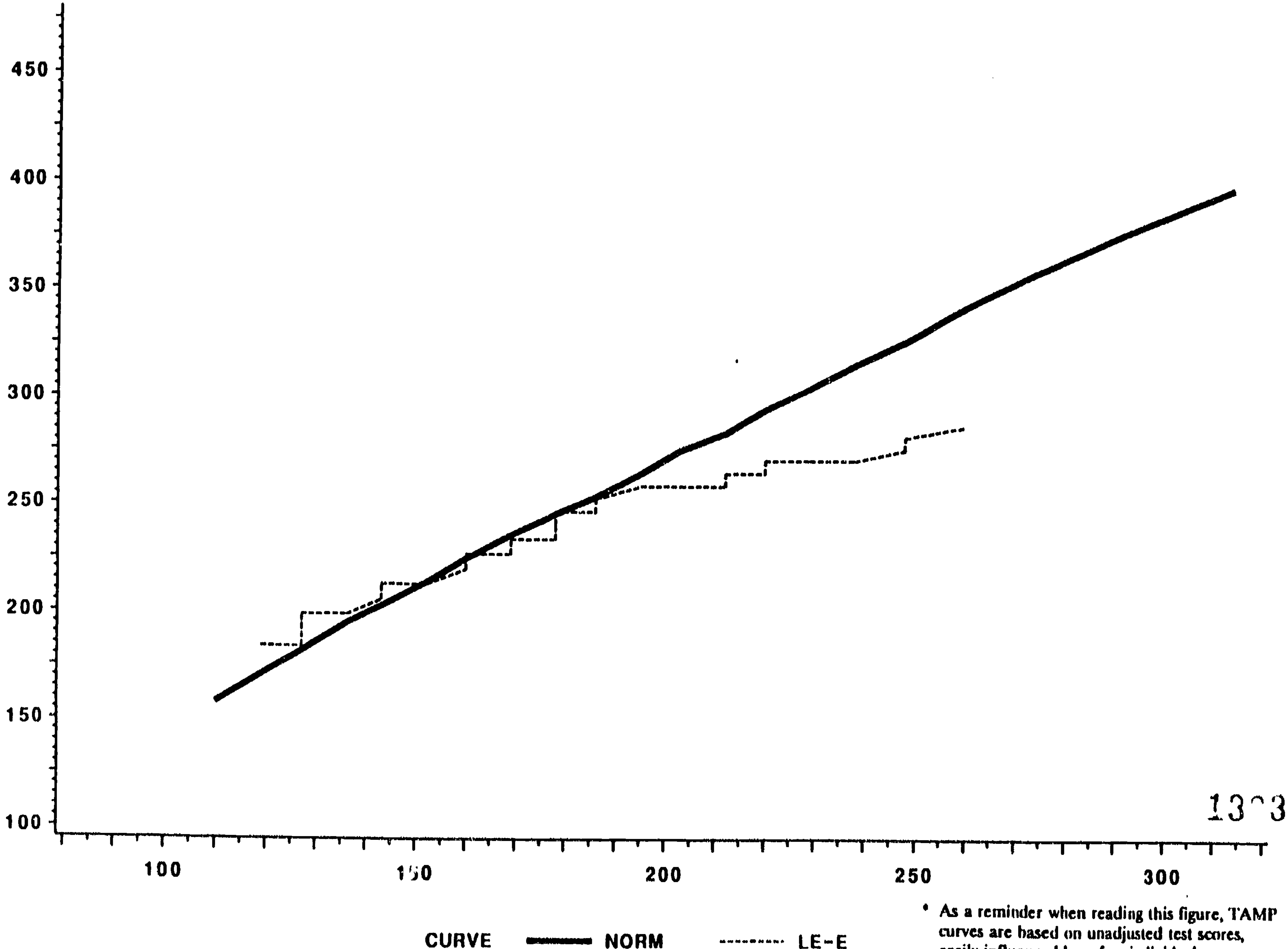


• As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases,

Figure 201

English Language Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



594

• As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

C.	**Grade Span:**	**Kindergarten**	**to**	**First Grade**
	Test Date:	**Fall**	**to**	**Spring**
	Language:	**Spanish**	**to**	**English**
	Content:	**Language**	**to**	**Reading**

Site D students with lower initial Spanish language skills learned to read in English more slowly than, whereas students with higher Spanish language skills learned to read as fast as, this norming population (see Figure 202). Site E students seemed to have a pattern of growth that was opposite that of site D students. Site E students with lower initial Spanish skills appeared to learn to read in English at about the same rate as this norming population while those with higher initial skills grew more slowly (see Figure 203). As with mathematics, this difference between the two sites for students at the upper levels of Spanish language skills may reflect the higher proportion of English instruction provided to site D students, and/or their more supportive home (with regard to socioeconomic and educational resources) and community environment.

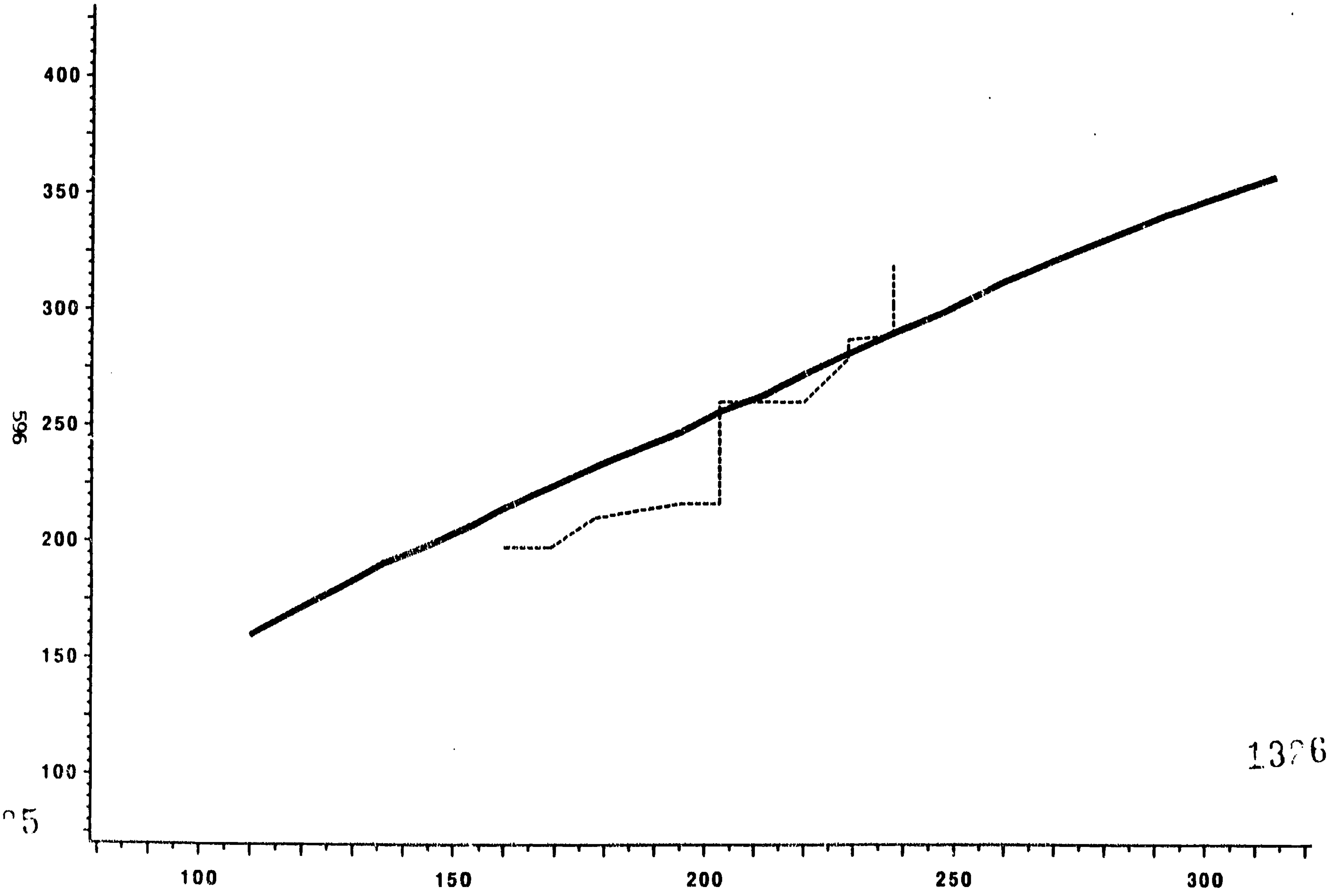
From another perspective, the difference between site D students who entered with lower primary language skills from those who entered with higher primary language skills would seem to support the notion that students with higher primary language skills would tend to acquire English reading skills faster when provided instruction in English than students with lower primary language skills.

In sum, while the lower English reading scores of site E students is consistent with the late-exit model predictions, the higher than expected performance of site D students suggests the potential effect of greater formal instruction in English as well as a variety of environmental factors supporting student learning. There is also some indication that students with higher primary language skills will tend to learn to read faster in English when instructed in English than students who have lower primary language skills.

Figure 202

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D

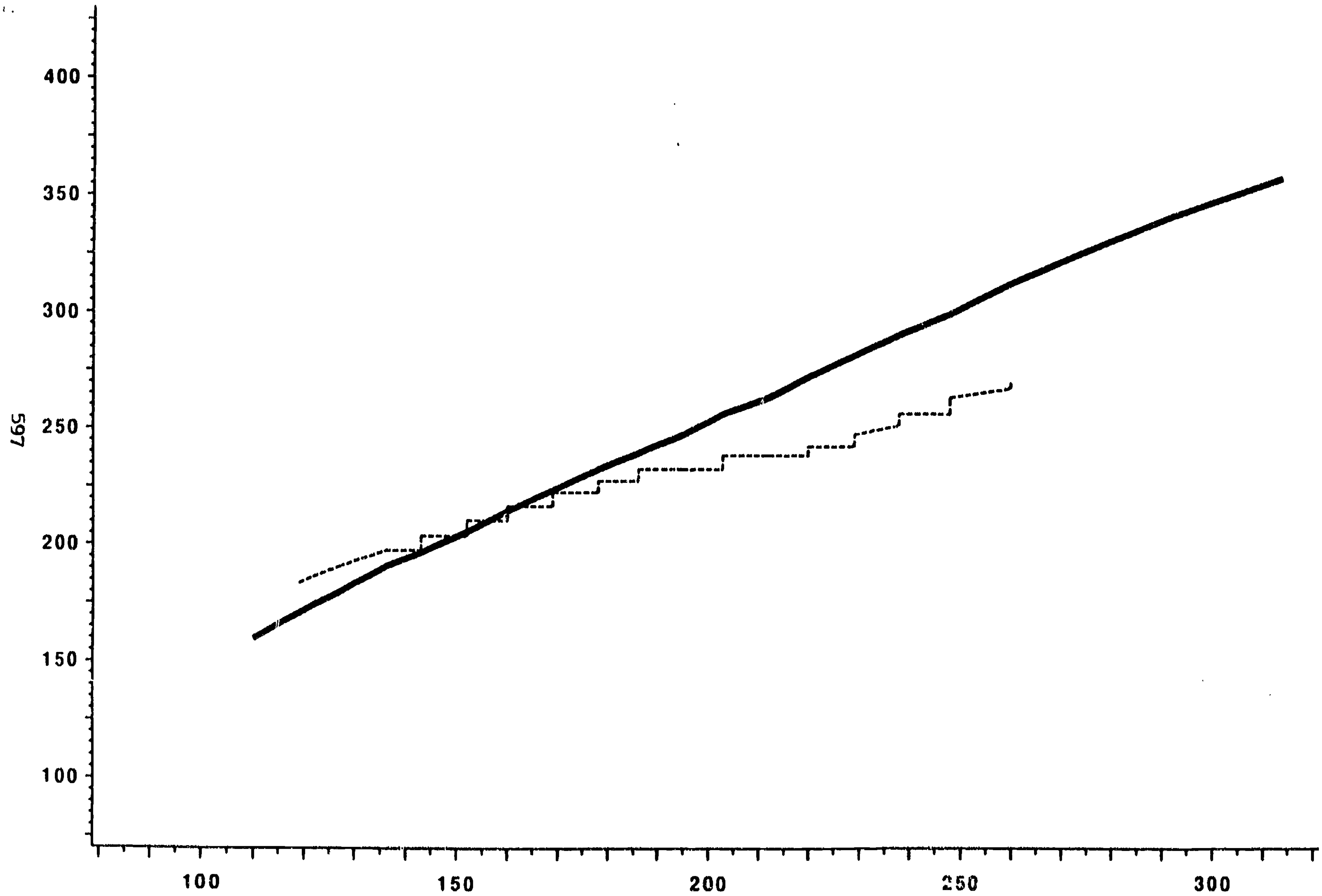


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 203

English Reading Spring 1 vs. Spanish Language Fall K (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, can be influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

4. First Grade to Third Grade:

A.	**Grade Span:**	**First Grade**	**to**	**Third Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

In mathematics, site D students grew faster, site E students at the same rate as, and site G students grew more slowly than this norming population (see Figures 204 to 206). This suggests that, if these trends were sustained over time, the average achievement level for site D students would be commensurate with that of this norming population and site E students would keep pace with this norming population. Site G students on the other hand would continue to fall behind relative to this norming population. Site E students with higher first grade scores seemed to learn at a faster rate than this norming population, while site E students with lower first grade scores seem to have learned at a slightly slower rate than this norming population. Once again, this was expected given that students received most of their instruction in Spanish, yet were assessed in English.

The distribution of growth rates for site D students indicates that these students began with higher first grade scores and ended with higher third grade scores relative to this norming population than did site E students. Compared to site D students, site E students began first grade and ended third grade with lower mathematics skills, presumably reflecting the more limited resources of family and community in site E.

Note that site G students began with almost as wide a range of first grade mathematics skills as site D and E students combined. Whereas site D and site E students consistently seemed to grow at a faster rate, or at least as fast as this norming population, site G students appear to have lost considerable ground consistently relative to this norming population. These differences are critical and appear to reflect the abrupt transition from Spanish to English instruction experienced in site G classrooms.

Figure 207 is a plot of the TAMP curves for each site relative to this norming population, thereby facilitating a comparison of each site.

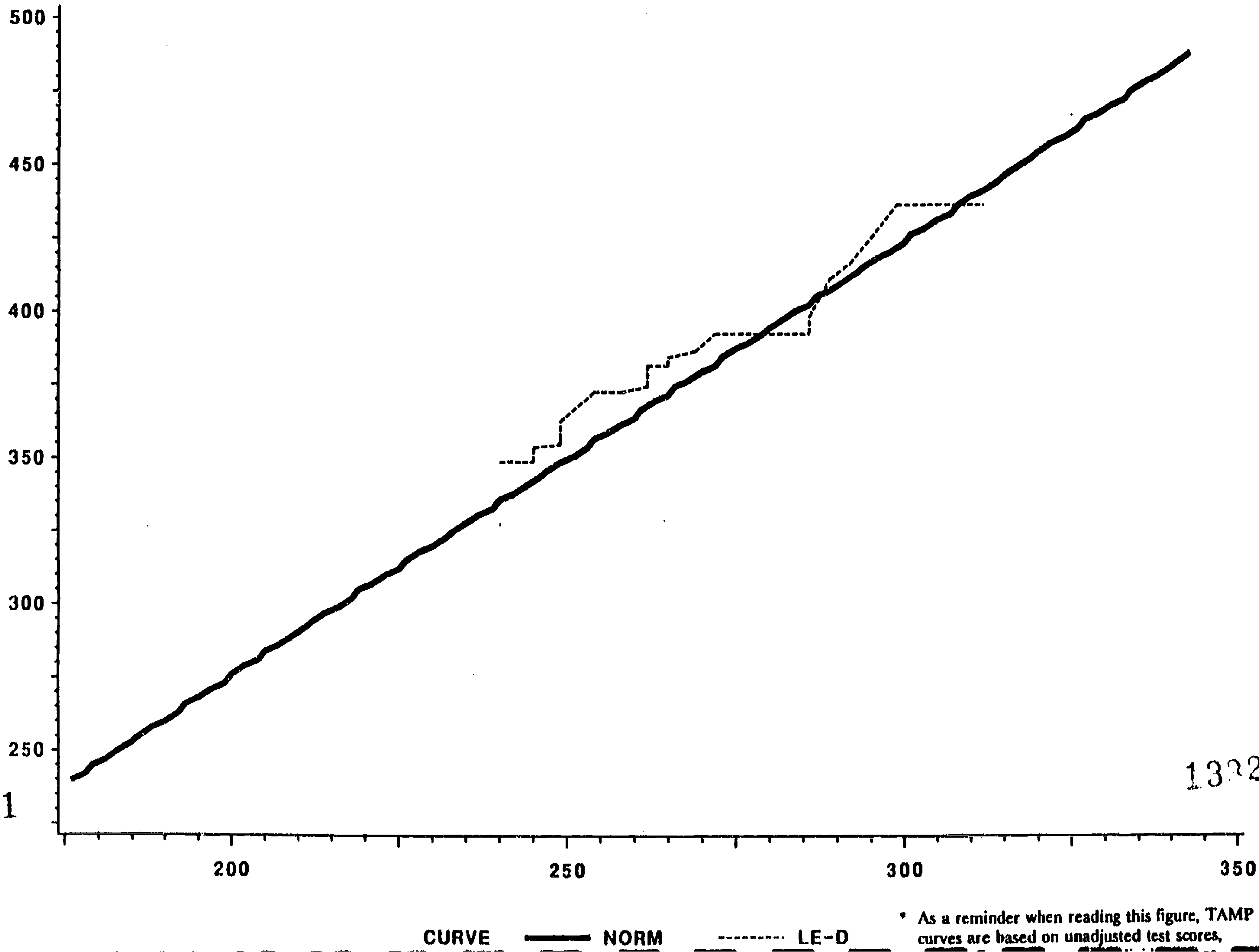
Recall that with respect to all major student, family and community characteristics save one, site G students, family, and community would lie somewhere between site D (with more resources) and site E (with the least resources). The exception is that in site G, while the use of primary language in grades one and two was approximately forty percent, it dropped abruptly to about one-fourth in grade three. In contrast, students in site D and E were receiving approximately forty percent or more of their instruction in Spanish through grade three.

In sum, while site G did not represent the site with the most resources nor the least, site G students ended first grade with mathematics skills comparable to that of site D and E students. Up through the end of first grade site G students had grown in their mathematics skills at a faster rate than site E students relative to this norming population. Yet site G students across the board lost ground from first to third grade relative to this norming population. This difference in growth rate is possibly attributable to the abrupt transition from primary language instruction into English instruction, which clearly seemed to have had a profound negative effect upon the learning of mathematics by site G students.

Figure 204

Math Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D

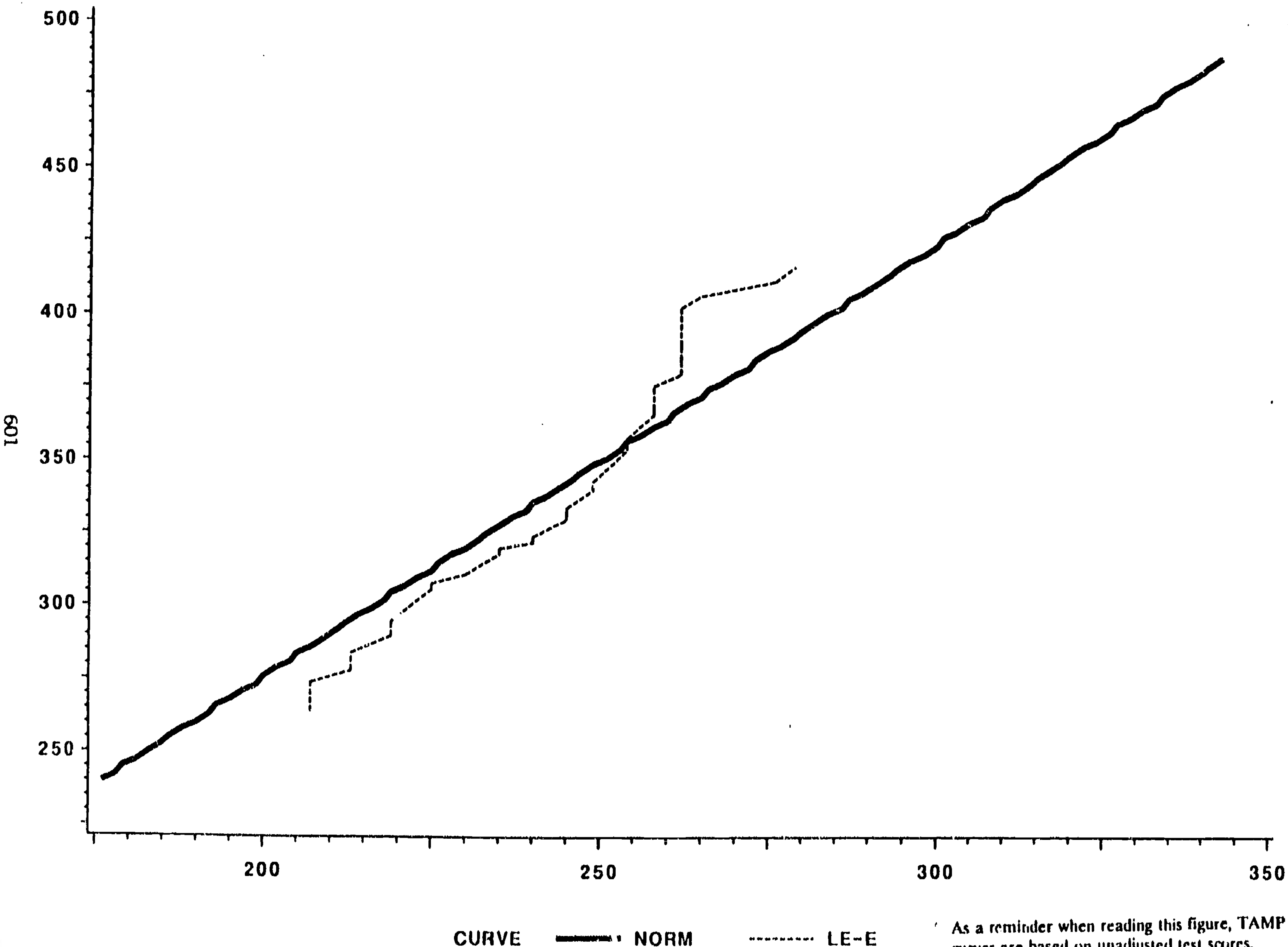


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores,

Figure 205

Math Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E



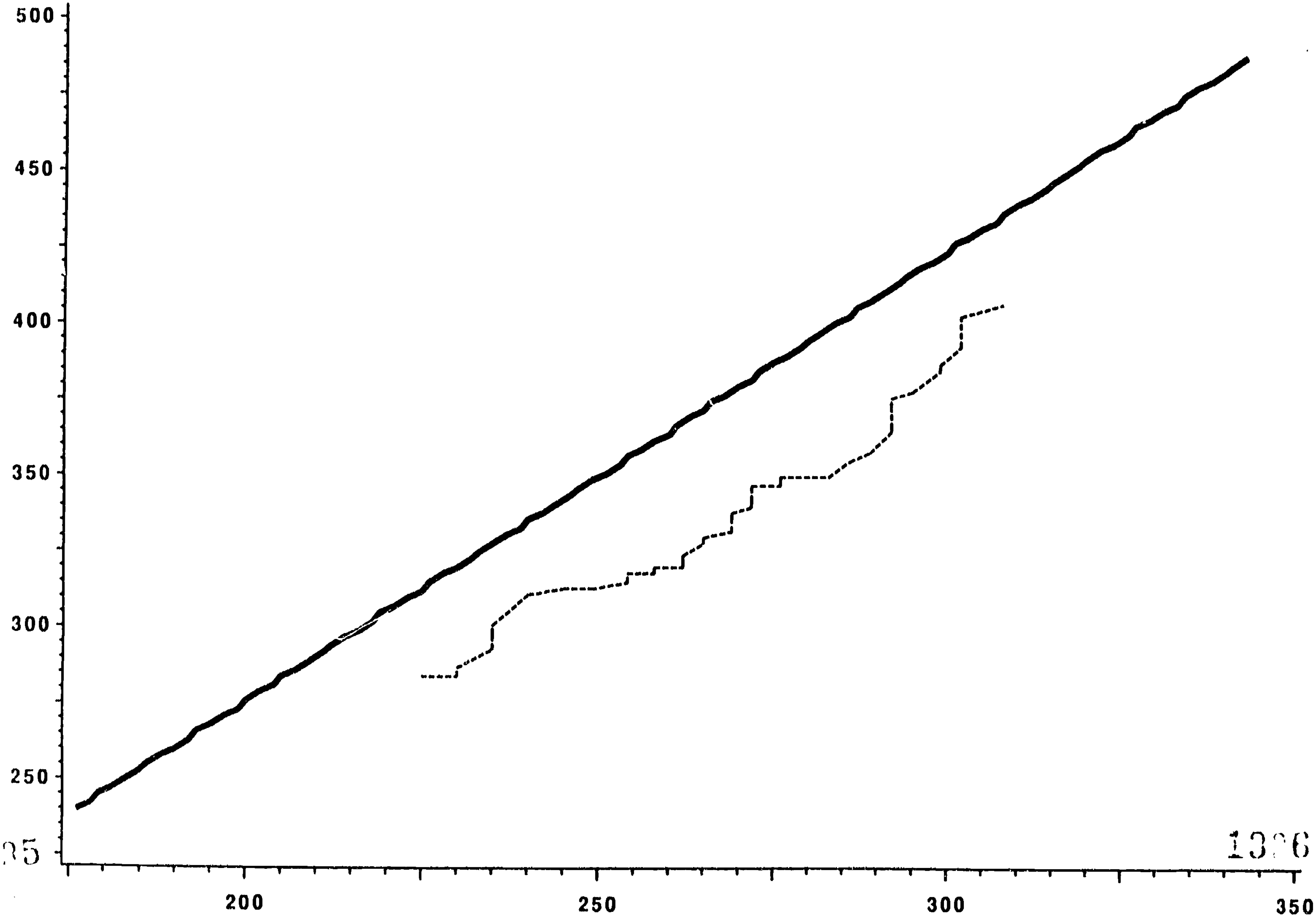
As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 206

Math Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G

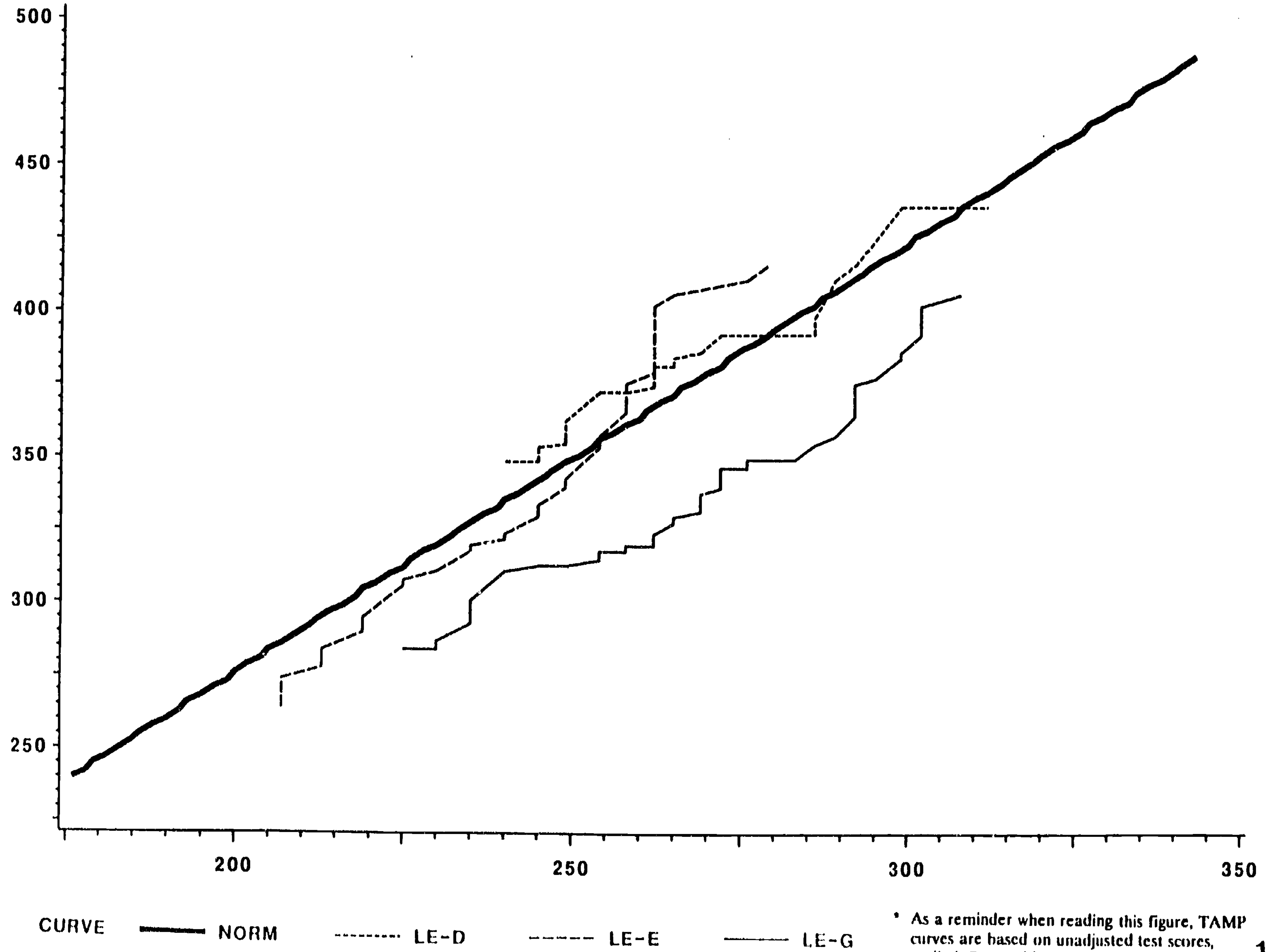


• As a reminder when reading this figure, TAMP

Figure 207

Math Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in Districts D, E, & G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation

B.	**Grade Span:**	**First Grade**	**to**	**Third Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Language**	**to**	**Language**

Figures 208 to 210 suggest that students in all three groups appear to be improving their English language skills at roughly the same rate relative to this norming population. Nonetheless, minor differences are evident between the sites.

Site D students began first grade with the highest distribution of first grade skills in comparison to students at sites E and G. Site D students with lower and higher skills grew as fast as this norming population, while those with average scores grew more slowly relative to this norming population. With so few students, the differences cannot be considered large. These differences within site D are consistent with the late-exit instructional model. It is assumed that as instruction in English is more limited, the more basic skills would be acquired more quickly while the more advanced skills would be acquired more slowly; hence greater growth of students with the lowest skills, and the slowest growth for those with higher skills.

Site E students with lower first grade scores may have been losing ground relative to this norming population (see Figure 209). The consistent growth rate among site E students with average or better first grade scores in English language skills relative to this norming population may reflect stronger primary and English language curricula in site E than in site D. Site E students had a stronger primary language instructional program than site D students in that site E students received more instruction in Spanish. Given that site E students with higher skills in first grade continued to grow at the same rate as this norming population speaks to the strength of the English curriculum to meet the needs of students at the middle and upper skill levels. The lower growth exhibited by site E students with lower first grade scores may either suggest a need to bolster the English curriculum at this level, or it may simply reflect the effect of a natural language approach. That is, students are encouraged to initiate use of English language skills without concern for the mechanical aspects of language (e.g., correct grammar, complete sentences, etc.). It would be expected that formal instruction in English is just beginning to be introduced into the third grade. As a result, students at these lower grades would not be pushed into the formal aspects of English language learning as fast as their classmates with stronger English language skills. Consequently, those students with lower initial scores would not be expected to grow at the same rate as this norming population.

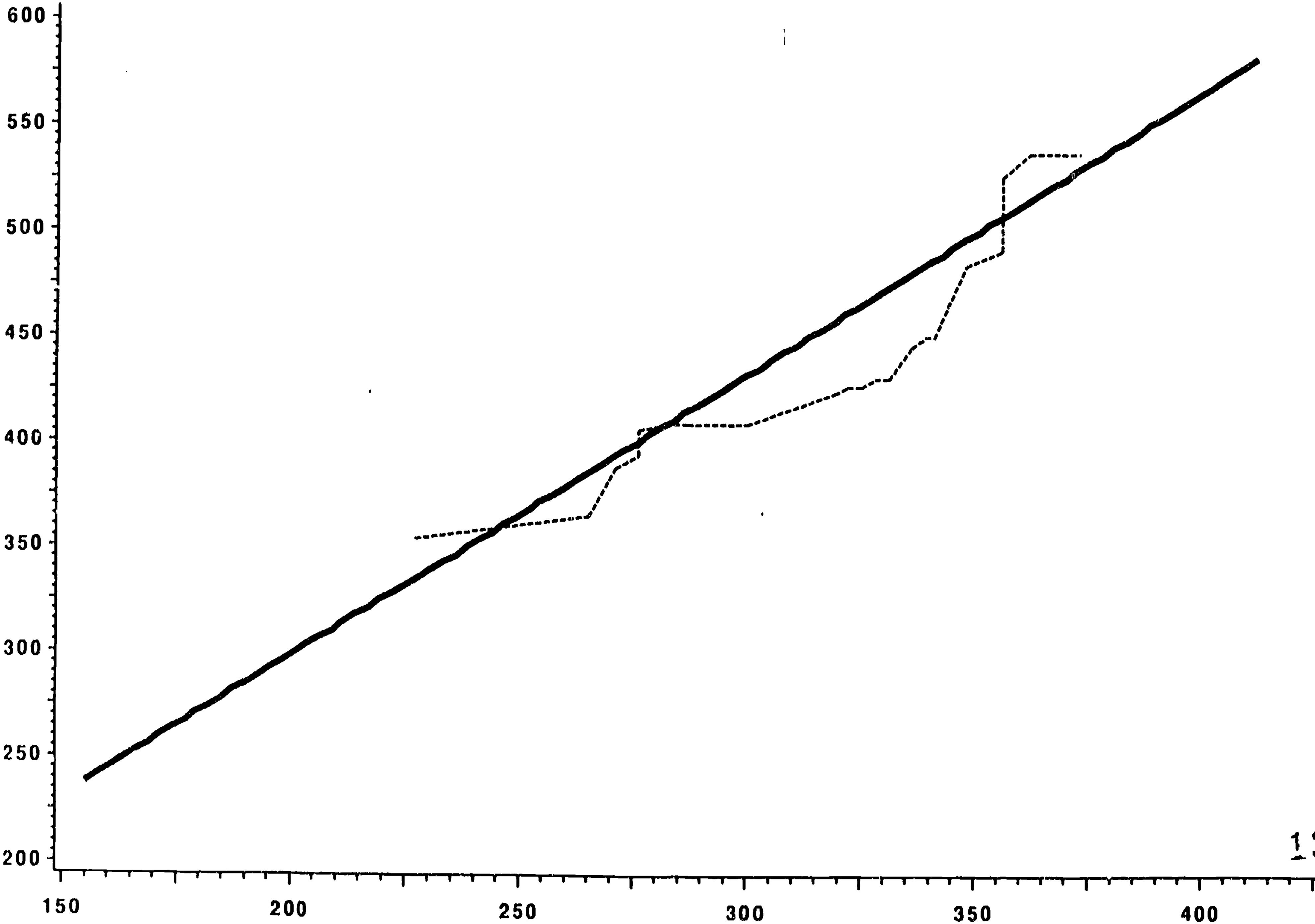
Site G students with average or below-average first grade language skills seemed to progress at the same rate as this norming population, but the students with higher first grade scores appear to have lost ground (see Figure 210).

These observations suggest that the English curriculum seems to address the needs of students within the average range. The instructional program does not appear to adequately meet the needs of students with higher entry-level English language skills. This slower growth for students with higher initial English skills is somewhat confusing given the abrupt increase in the amount of English instruction provided to site G students. Common sense would suggest that of all students, those with higher English skills would do better with greater instruction in English. Once again, this slower growth exhibited may reflect, albeit at a milder level, the effect of the abrupt increase in English instruction and/or a failure in the English curriculum to meet the needs of students with higher English skills.

Figure 208

Language Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



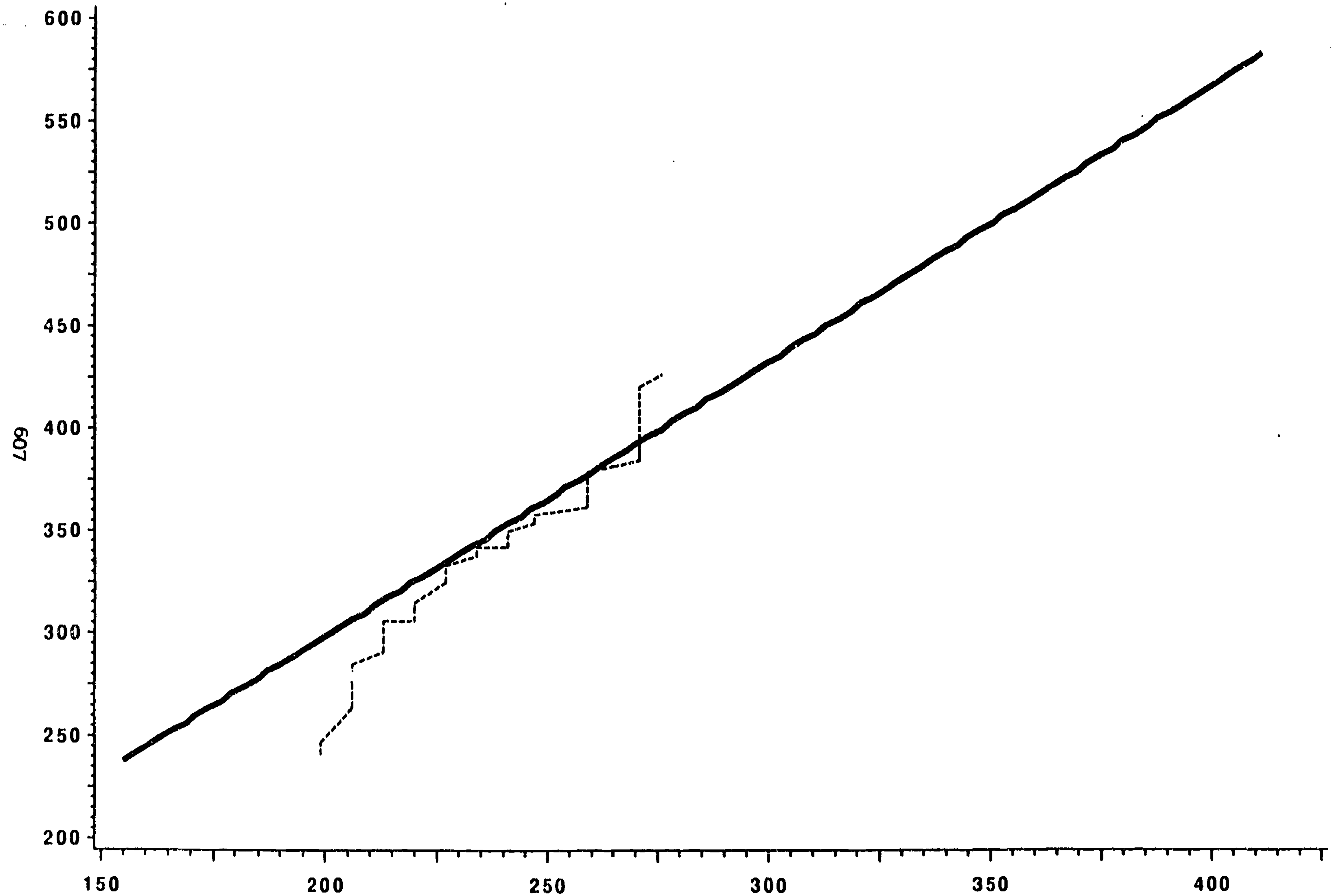
CURVE —— NORM -------- LE-D

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 209

Language Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E

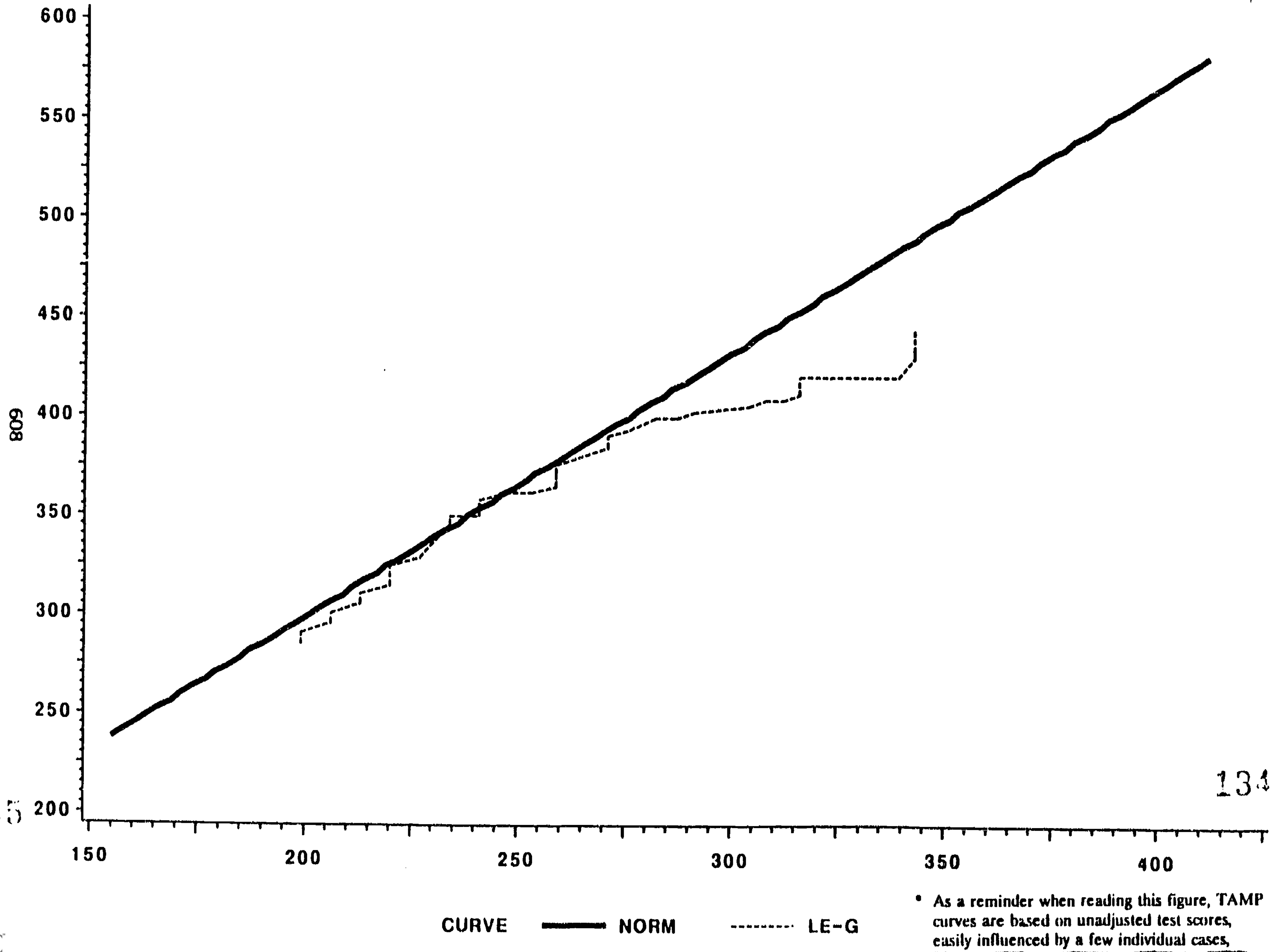


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 210

Language Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

1345

1346

ERIC

C.	**Grade Span:**	**First Grade**	**to**	**Third Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Reading**	**to**	**Reading**

Overall, site D and site E students seemed to be increasing their English reading skills at the same rate or faster than this norming population (see Figures 211 and 212). Although site G students spanned the same range of skills as site D and site E students taken together, across all initial (first grade) reading levels site G students seemed to be losing ground relative to this norming population (see Figure 213). Once again, although site G students did not have the strongest nor the weakest resources in their families or communities and began with roughly the same skills as students in the other two sites, their growth in English reading was the slowest and of most concern.

Site D students with lower first grade reading scores seemed to be growing at the same rate as this norming population. The dip in reading growth among a small section of site D students with higher reading scores may reflect random fluctuation, insofar as those with the highest entry-level reading scores were growing at the same rate or slightly faster relative to this norming population.

Site E students, especially those with the higher first grade scores, appeared to grow at a faster rate relative to this norming population. This is surprising given that site E students had less English instruction than either site D or site G students. Site E students with the lowest initial scores grew about as fast as this norming population.

Students with less English instruction were projected to grow at a much slower rate than students receiving greater amounts of English instruction. This is clearly not borne out by the data when site E (having the least instruction in Spanish of the late-exit sites, students with the lowest distribution of initial English reading scores, and a program context where family, school and community resources are extremely

ERIC

limited) is compared to site G (having the most instruction in English among the late-exit sites, proportionately more students with higher English reading skills, and better family, school and community resources).

TAMP analyses indicate that site G students are losing ground relative to this norming population. As with mathematics, it is proposed that this slower growth rate reflects the effect of the abrupt transition into English reading. It may also reflect, in part, a less than ideal English reading program for limited-English-proficient students.

To summarize, the TAMP analyses for grades one to three were effected to better understand the findings presented in Chapter V which compared the performance of the three late-exit sites. The TAMP analyses also provide us with a way to examine more closely the unexpected results presented earlier in the first section of this chapter, i.e., that late-exit students seemed to be growing as fast as or faster than this norming population in mathematics, language arts, and reading when tested in English.

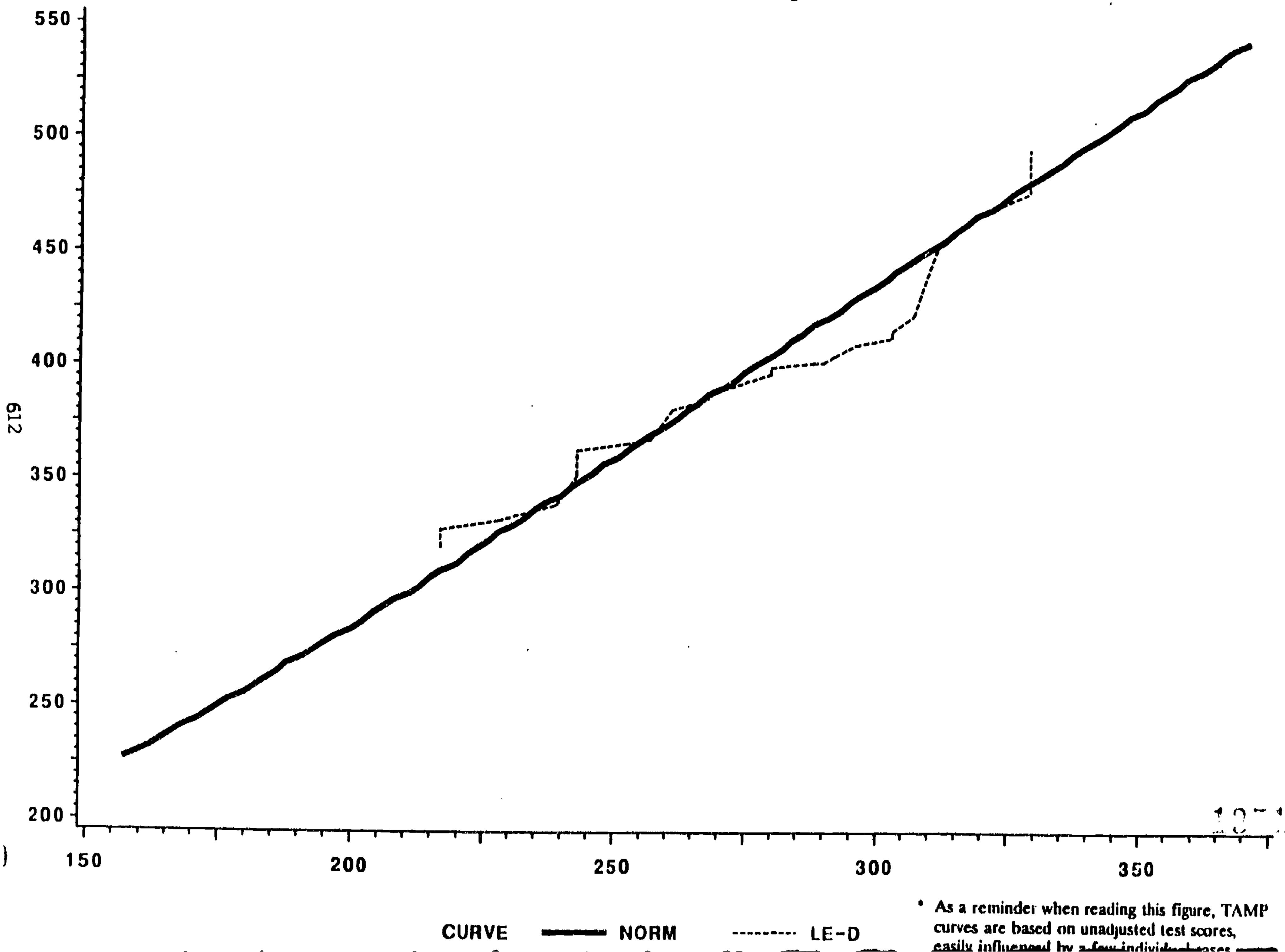
From the late-exit instructional model we predicted that the growth rate of skills among late-exit students would be slower not only relative to this norming population, but to that of immersion strategy and early-exit students. That late-exit students often seemed to show growth rates almost as high as or even better than immersion strategy students, early-exit students, and this norming population was not expected for this grade range. Typically the performance of language-minority students in the general population shows an increased decline in their academic growth over time. Minimally the TAMP analyses support the need for and effectiveness of special services for limited-English-proficient students. The first to third grade TAMP analyses seem to suggest two findings. First, the analyses seem to firmly support the hypothesis that students who receive more instruction in their primary language and allow for a more gradual introduction in instruction in English will acquire skills in a second language faster than students who are abruptly transitioned into instruction in English and less in Spanish. Secondly, the TAMP analyses seem to document the negative effects upon students' learning rates when there is an abrupt shift from primary language instruction to instruction in English.

ERIC Full Text Provided by ERIC

Figure 211

Reading Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District D



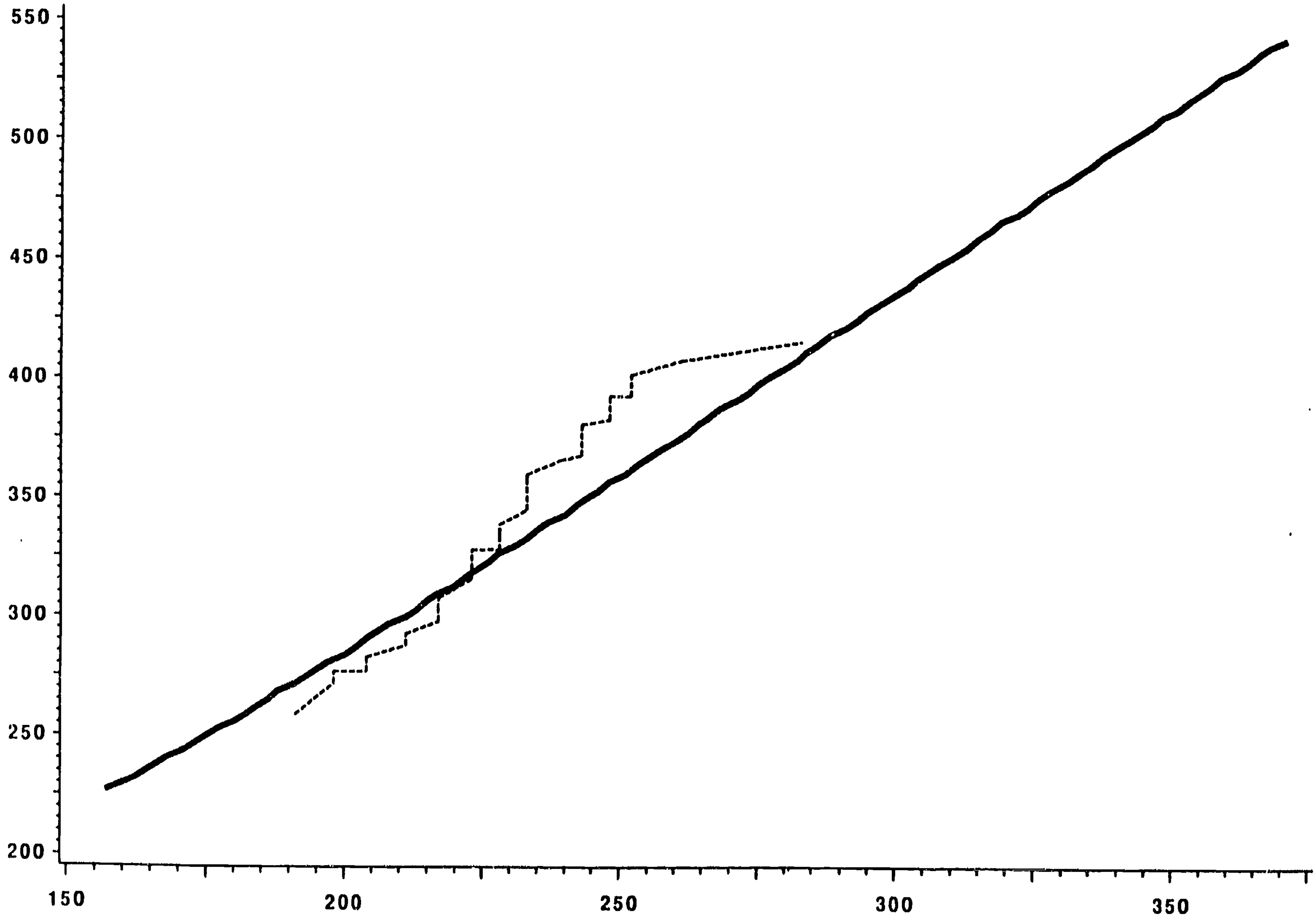
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases

ERIC

Figure 212

Reading Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E

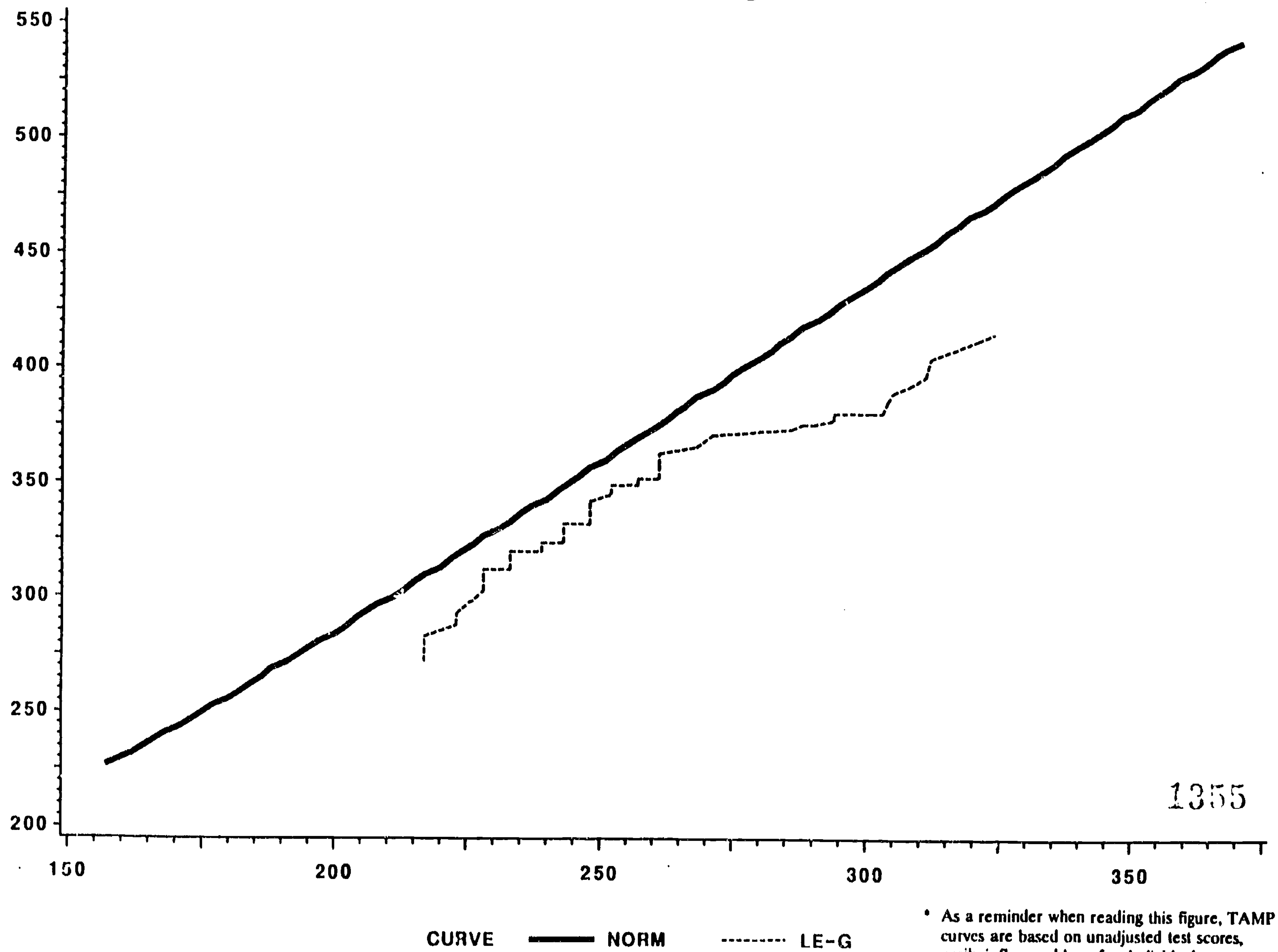


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 213

Reading Grade 3 vs. Grade 1 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, [illegible]

5. Third Grade to Sixth Grade:

A.	**Grade Span:**	**Third Grade**	**to**	**Sixth Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

Figure 214 is a TAMP curve created by collapsing across the three late-exit sites. This combined analysis suggests that late-exit students with lower third grade scores appear to have been learning at a slightly slower rate, yet those with higher scores were learning at the same rate or also slightly slower relative to this norming population. However, when the TAMP curves for each late-exit site are compared (see Figure 215), the between-site differences in growth rates in mathematics is striking. Of the three sites, students of average or higher initial mathematics scores at site E (the site with the lowest distribution of third grade scores, the fewest resources, the most instruction in Spanish, and the least instruction in English) consistently seemed to learn at the same rate relative to this norming population and to comparable site D or site G students. Site E students with lower initial mathematics scores appear to have lost ground when tested in English. Moreover, and most importantly, the growth rate by site E students is consistent with the underlying hypothesis of the developmental primary language program. That is, providing content instruction in the primary language ensures that language-minority children do not fall behind, but continue to increase their skills. If this observed growth rate were sustained, over time the distribution of achievement scores site E students would be similar to this norming population.

Students in site D (with the highest distribution of third grade scores, most resources, second greatest amount of Spanish instruction, and the second least amount of English instruction) with average third grade scores seemed to grow at about the same rate as this norming population. Students with the highest and lowest third grade scores appear to have lost ground relative to this norming population. This fluctuation may reflect the limited number of cases at this end of the distribution. Despite the uneven pattern of growth, on average site D students with low and average initial skills appear to be tracking the growth rate for this norming population. Once again, this finding is atypical for these students and seems to be supportive of the use of primary language instruction. Perhaps the slower and more uneven growth rates as compared to those for site E students may reflect the lower use of Spanish in site D than in site E classrooms.

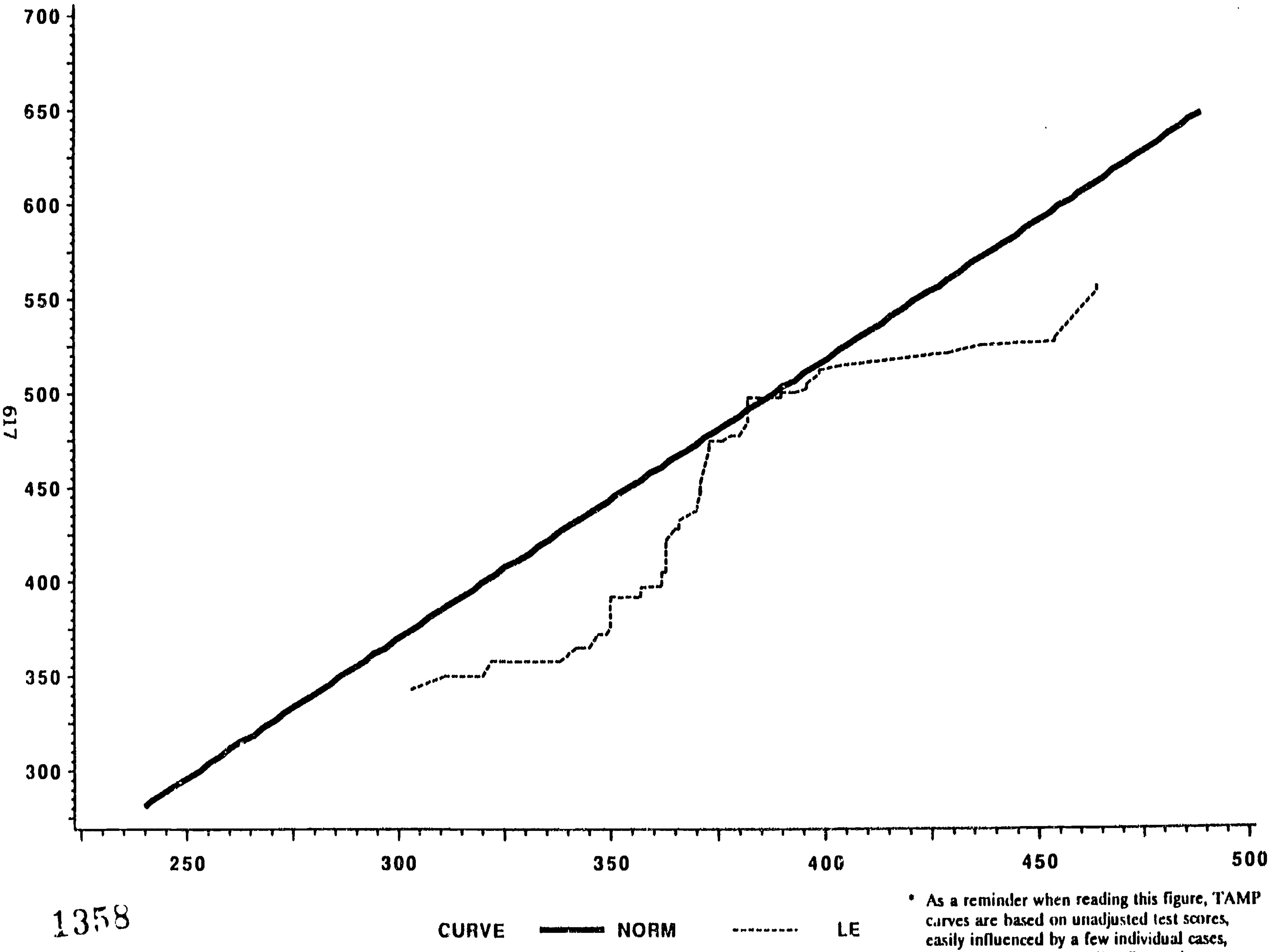
Once again, Figure 215 seems to illustrate clearly the negative effects upon the acquisition of mathematics of abruptly switching from primary language instruction to English instruction. Site G students regardless of their initial third grade scores seemed to consistently lose ground relative to this norming population. Notice that they already appear to be growing more slowly than they did from first grade to third grade.

In sum, site E students seemed to improve their skills in mathematics, followed by site D students, at a rate close to that of this norming population. Site G students not only lost ground, but the entire distribution of scores narrowed and shifted downward relative to this norming population. Once again, this may indicate that shifting abruptly from Spanish to English has a strong negative impact upon the rate of academic growth of limited-English-proficient students.

Figure 214

Math Grade 6 vs. Grade 3 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



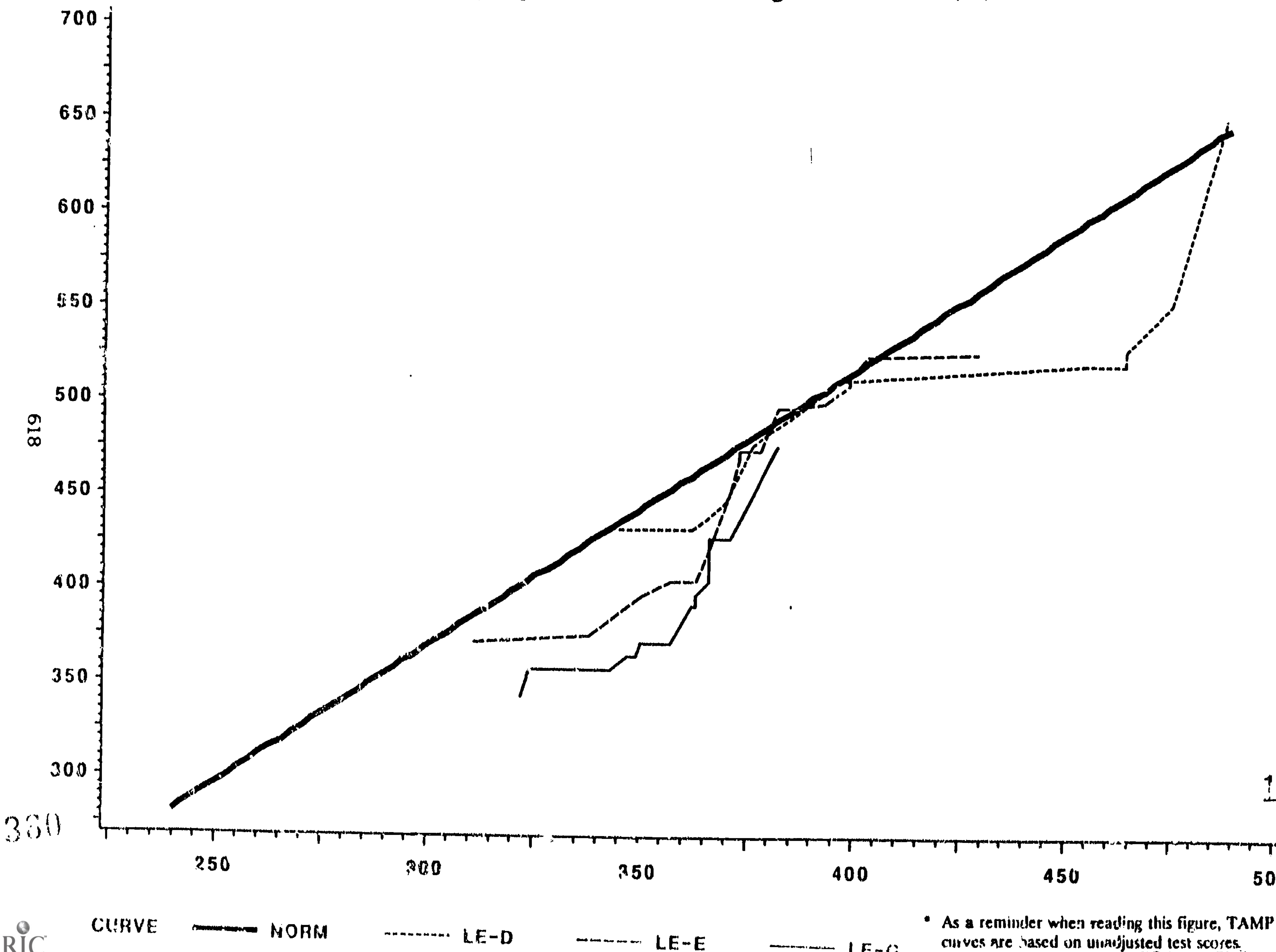
- As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

Figure 215

Math Grade 6 vs. Grade 3 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in Districts D, E, & G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores.

B.	Grade Span:	Third Grade	to	Sixth Grade
	Test Date:	Spring	to	Spring
	Language:	English	to	English
	Content:	Language	to	Language

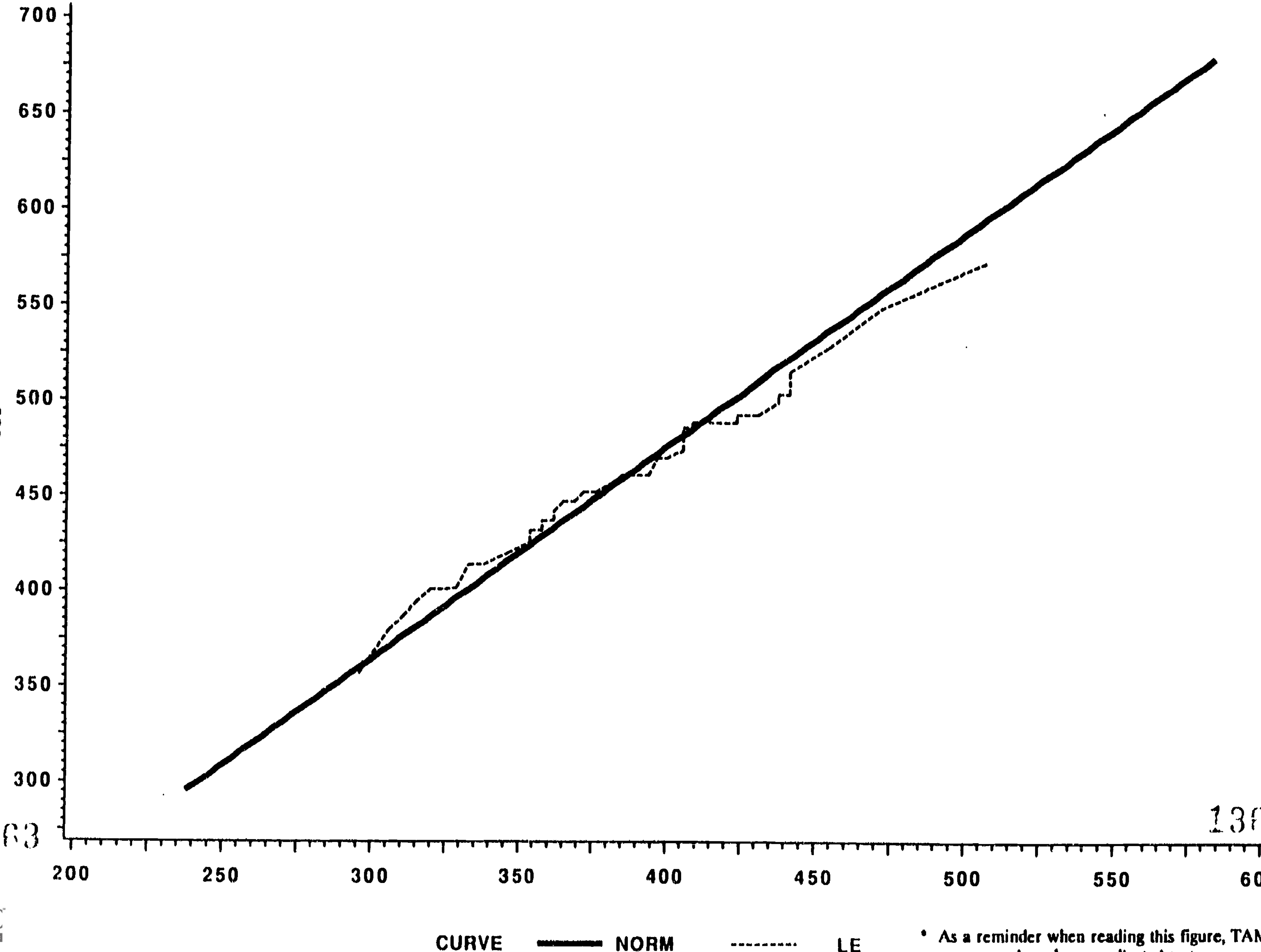
When scores are combined across the three late-exit sites (see Figure 216), it seems that the average rate of growth is approximately the same as that of this norming population. As with mathematics, when the TAMP curves are plotted for each site against the growth curve of this norming population there are marked differences between sites (see Figure 217). The differences between sites in the distribution of English language skills at third grade are similar to those observed for mathematics.

Regardless of their original third grade scores, students in sites D and E consistently learned at about the same rate as this norming population. While site G students with the lowest and average scores appeared to grow at about the same rate as this norming population, those with higher scores may have lost ground. As with mathematics, these findings seem to support the notion that the development of primary language skills facilitates the rate at which second language skills are acquired.

Figure 216

Language Grade 6 vs. Grade 3 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



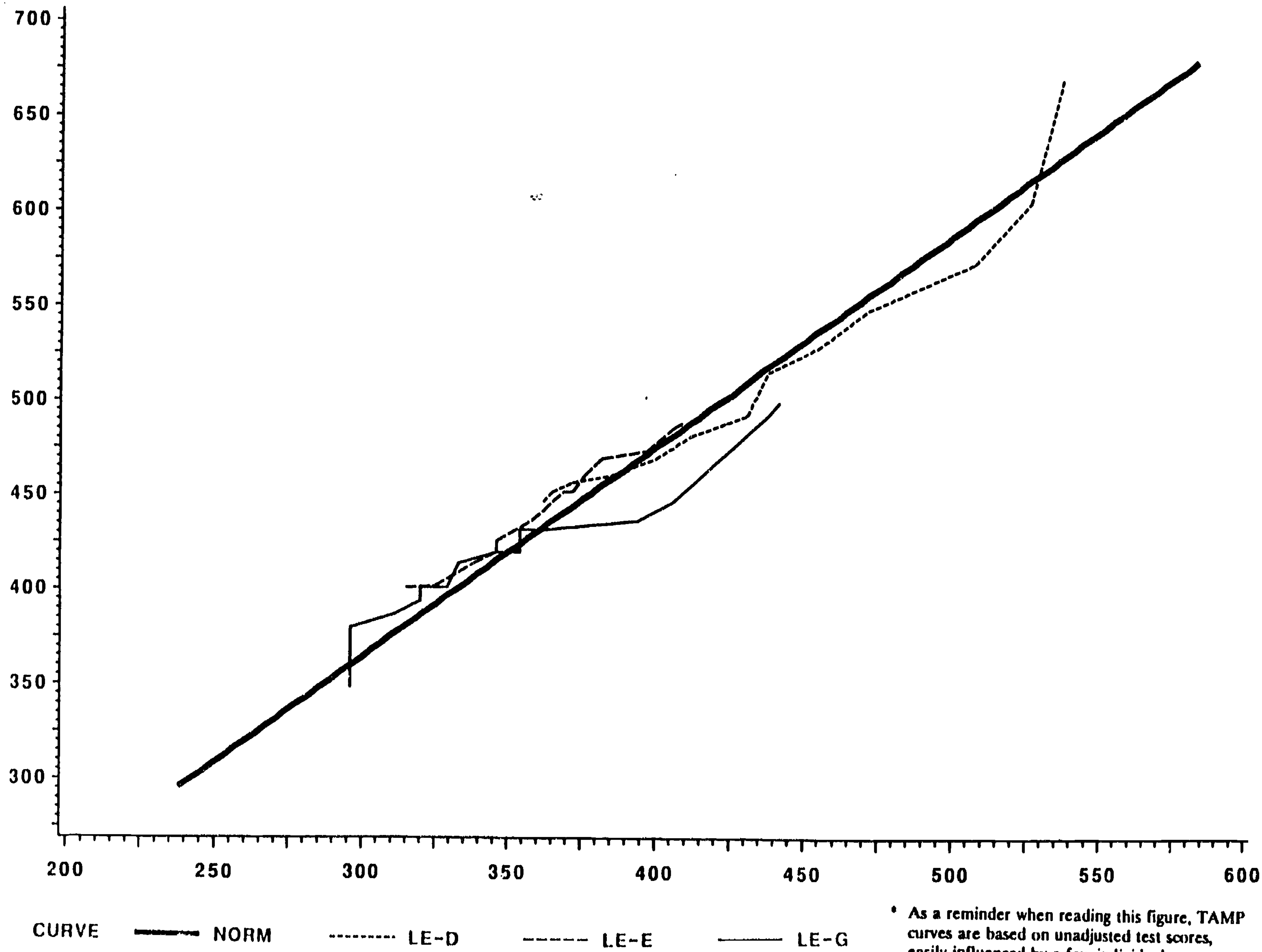
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores.

ERIC

Figure 217

Language Grade 6 vs. Grade 3 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in Districts D, E, & G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

C.	**Grade Span:**	**Third Grade**	to	**Sixth Grade**
	Test Date:	**Spring**	to	**Spring**
	Language:	**English**	to	**English**
	Content:	**Reading**	to	**Reading**

TAMP analyses of the English reading scores from across all late-exit sites suggest that all students appear to have grown at the same rate, which was slightly slower than this norming population (see Figure 218). Late-exit students with the highest initial reading scores seemed to be losing the most ground. As in the case of mathematics and language, marked differences are noted when the TAMP curves of each site are plotted against this norming population (see Figure 219).

In sum, it appears that site D and site G students with the lowest initial English reading skills grew at the same rate as this norming population. Those site D and site G students with higher initial skills grew the least. It seems that site E students grew slower than this norming population.

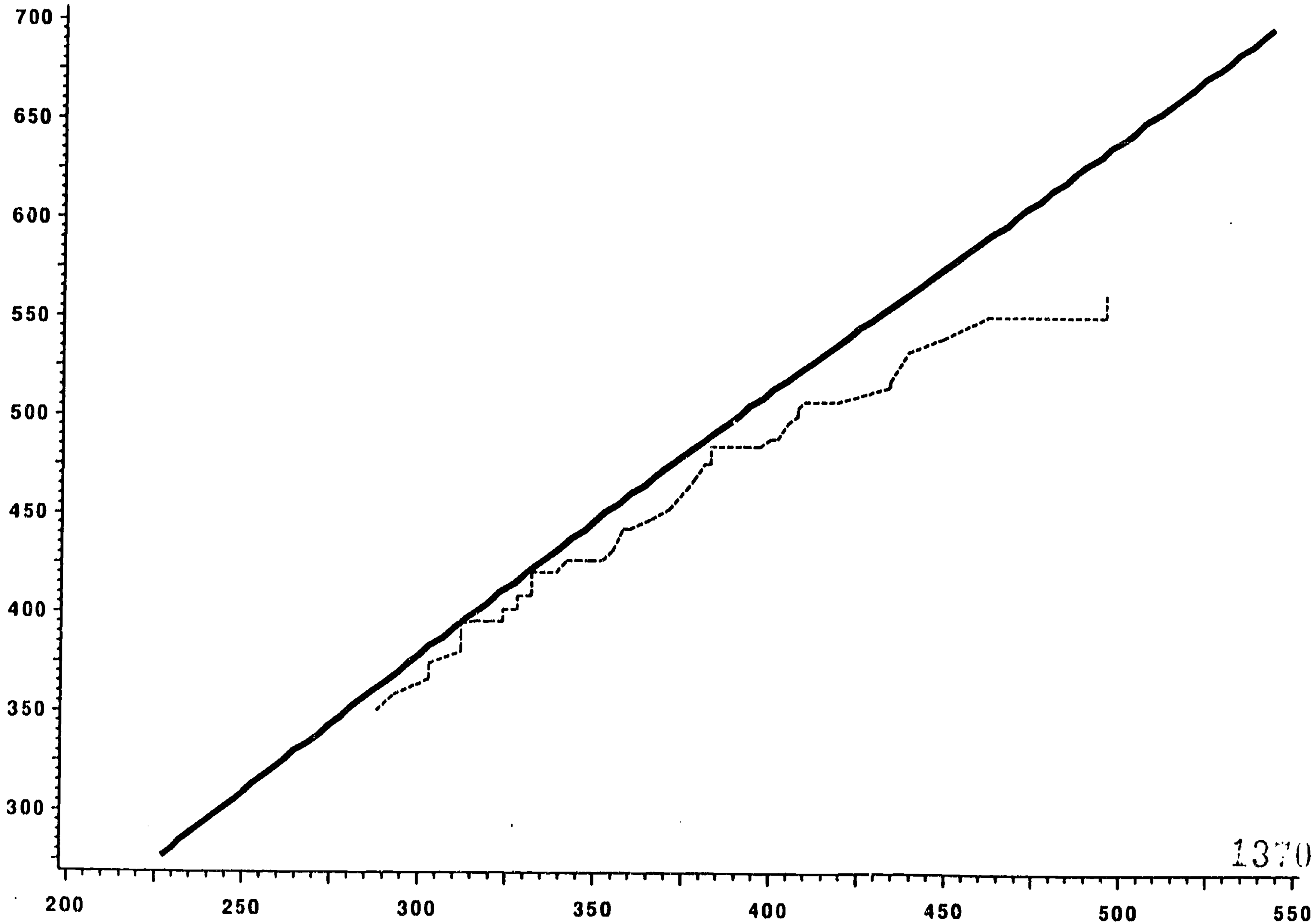
To summarize, the third to sixth grade cohort analyses of the growth in mathematics, English language, and English reading skills are consistent with the hypothesis of the facilitative effects of primary language development upon second language acquisition and academic achievement. Those students who had substantial amounts of instruction in Spanish and a slow increase in the use of English for instruction grew the fastest relative to this norming population. This growth was realized even though these students had the lowest distribution of initial academic skills and came from homes, schools, and a community with the fewest resources and the most risk. As the amount of primary language instruction decreased, student growth tended to be slower and more inconsistent among site D and site G students. Those site D and site G students receiving the least amount of primary language instruction and the most instruction in English consistently lost ground relative to this norming population. These findings are very supportive of the underlying hypothesis for primary language instruction. That is, limited-English-proficient students will acquire second language skills faster if they are given the opportunity to develop their primary language skills first.

ERIC

Figure 218

Reading Grade 6 vs. Grade 3 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



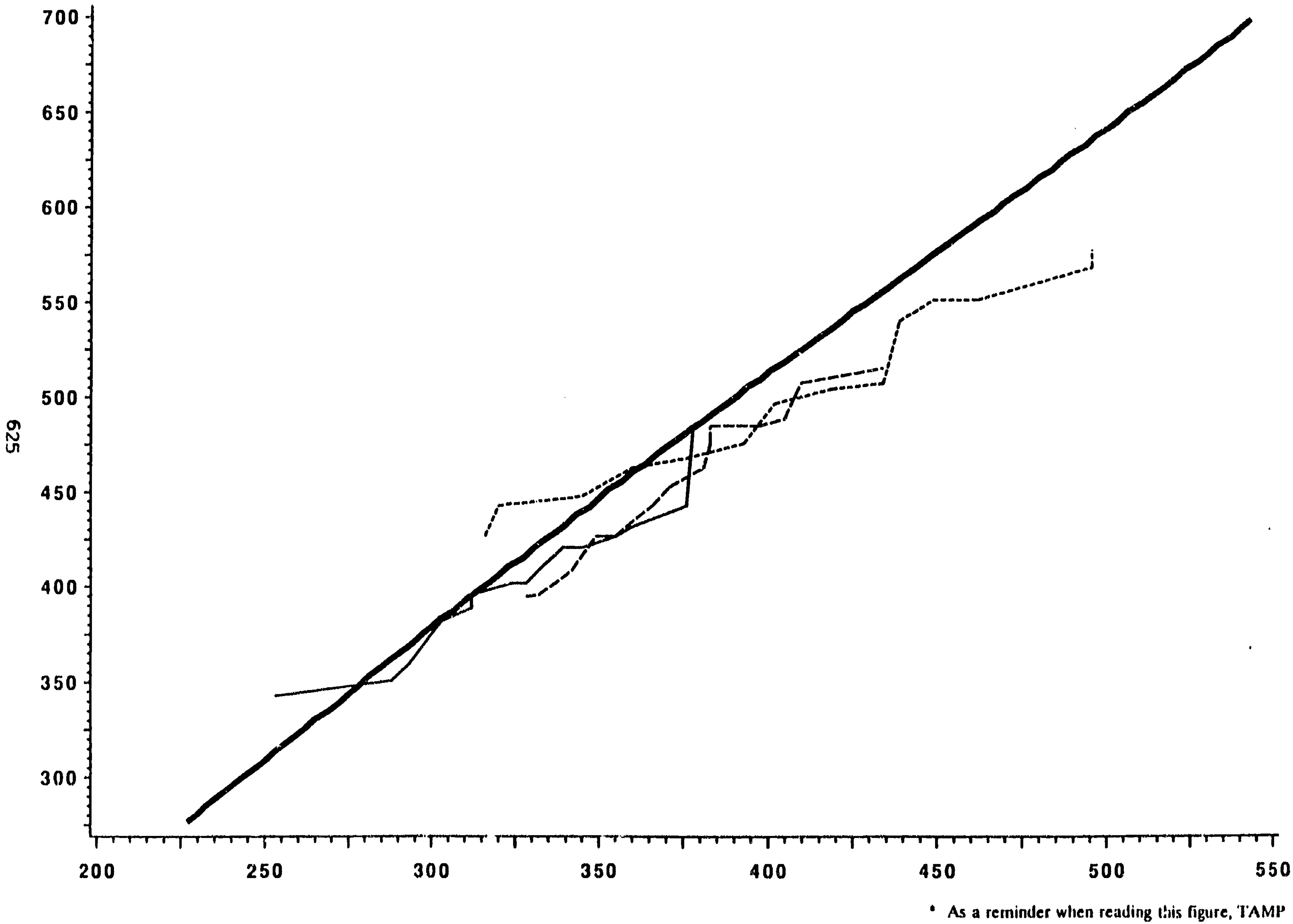
* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling

ERIC

Figure 219

Reading Grade 6 vs. Grade 3 (Matched Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in Districts D, E, & G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

ERIC

6. First Grade to Sixth Grade:

The first grade to sixth grade TAMP analyses are a composite of the data from the K-3 cohort with the 3-6 cohort. The assumption is made that the characteristics of students, their families, school, and community are essentially the same. Our examination of these contextual variables in Chapter V did not uncover any significant differences between the two cohorts. This combined analysis provides us with a view of the growth in academic achievement for late-exit students from first to sixth grade. This permits us to examine the TAMP curves to look for evidence that might confirm or reject the underlying hypothesis of the facilitative effects of primary language instruction for limited-English-proficient students. That is, is there evidence that a primary language development program will facilitate the acquisition of English language skills and skills in content areas? The late-exit model would predict low achievement from kindergarten through about the third grade and a slow increase in achievement beginning in grade four or five, eventually approximating the achievement of the norming population by grade six. This was the focus of discussion for Chapter V. In terms of rate of growth, to realize these achievement patterns limited-English-proficient students would need to demonstrate a growth rate that was greater than that of this norming population. This is necessary because the skill levels of limited-English-proficient students are lower than those for this norming population. Thus if limited-English-proficient students were to catch up with this norming population, they would have to be growing at a faster rate than this norming population. If this were realized the late-exit programs would clearly be successful in meeting the needs of limited-English-proficient students and in meeting the goals of these programs; that is, to ensure the development of skills in English and in the content areas for limited-English-proficient students to the point where they could successfully compete with English-only students.

From a slightly different perspective, the late-exit programs would also be deemed a success if limited-English-proficient students grew at the same rate as this norming population. This would imply that given the characteristics of limited-English-proficient students (e.g., socioeconomic status), if they were able to maintain the same growth rate (i.e., percentile standing) as their counterparts in this norming population, they would be doing as well as one could expect. Given that the growth rate for language-minority students in the general population continually decreases over time, we would expect an ever widening gap between their performance and that of this norming population. If the limited-English-proficient students in the late-exit programs were to grow at the same rate as this norming population, this would be evidence of the effectiveness of the late-exit programs and the instructional model.

A.	**Grade Span:**	**First Grade**	**to**	**Sixth Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Mathematics**	**to**	**Mathematics**

There is evidence to support the late-exit hypothesis that a strong primary language development program will help limited-English-proficient students to successfully acquire English language skills and content area skills. Figure 220 shows the TAMP curve for all late-exit students in mathematics collapsed across the three late-exit sites. It suggests that the distribution of initial first grade scores for the late-exit students is shifted slightly to the left, indicating lower initial skills among these students relative to this norming population. This figure also suggests that on the average, the growth rate for all late-exit students taken together seems to approximate the growth rate for this norming population. Upon closer inspection one notes that while the rate of growth for students with lower initial first grade scores was slightly lower than this norming population, the growth rate in mathematics for students with higher initial skills appeared to be faster than that of this norming population. The highest scoring students mirrored the growth rate of this norming population. All of these findings indicate a measure of the effectiveness of the late-exit programs in helping limited-English-proficient students to maintain their position relative to this norming population. If sustained over time, a faster rate of growth than this norming population would result in an upward shift in the achievement and position of these students relative to this norming population. However, there appear to be marked differences in the patterns of growth between the three late-exit sites.

Figure 221 strongly suggests that the differences in the growth of mathematics between late-exit sites D, E, and G may reflect the observed differences in the amount of primary language instruction provided. Site E (with the lowest distribution of scores and the fewest family, school, and community resources), which provided the most instruction in Spanish over time, had students at all initial skill levels grow at slightly to substantially faster rates relative to this norming population. If this growth rate were sustained, in time one would expect that site E students would catch up to the average achievement level of this norming population.

Site D (with more students with higher initial mathematics skills, and family, school, and community resources) provided its students with a consistent pattern of primary language instruction through the fifth grade. Students were provided with approximately forty percent of their instruction in Spanish. These students seemed to grow at the same rate as or slightly faster than this norming population, except that students with initial low mathematics skills lost ground relative to this norming population. Given the limited number of students at this site who completed the study, the points at the extreme ends of the curves might be somewhat anomalous.

Dramatically, site G students with the least amount of primary language instruction (but with students spanning the range of skills of the other two sites, and with family, school, and community resources falling between the other two sites) consistently lost ground relative to this norming population. Recall from Chapter I that the pattern of use of English and Spanish more closely resembled the early-exit model than the late-exit model for site G. Students were provided with credible amounts of instruction in Spanish (i.e., at least 40%) through the second grade.

Thereafter, almost all of their instruction was in English. This TAMP analysis suggests the profound negative impact upon the growth rate in mathematics skills for limited-English-proficient students when there is such an abrupt change in their language of instruction. The growth rate of these students continues to drop in percentile rank relative to this norming population. This relative decline parallels the observations of the achievement of disadvantaged students in the general population.

Figure 220

Math Grade 6 vs. Grade 1 (All Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program

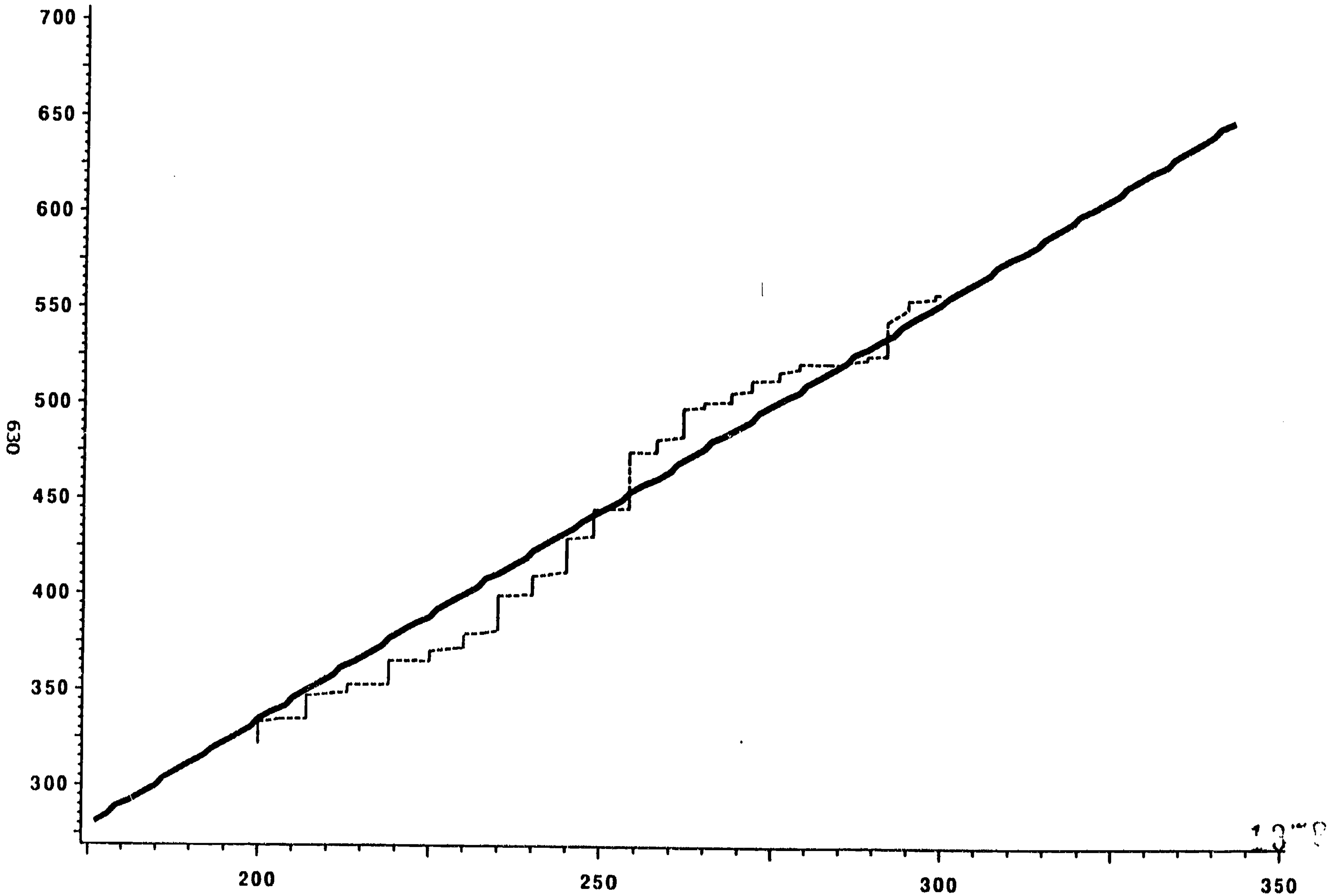
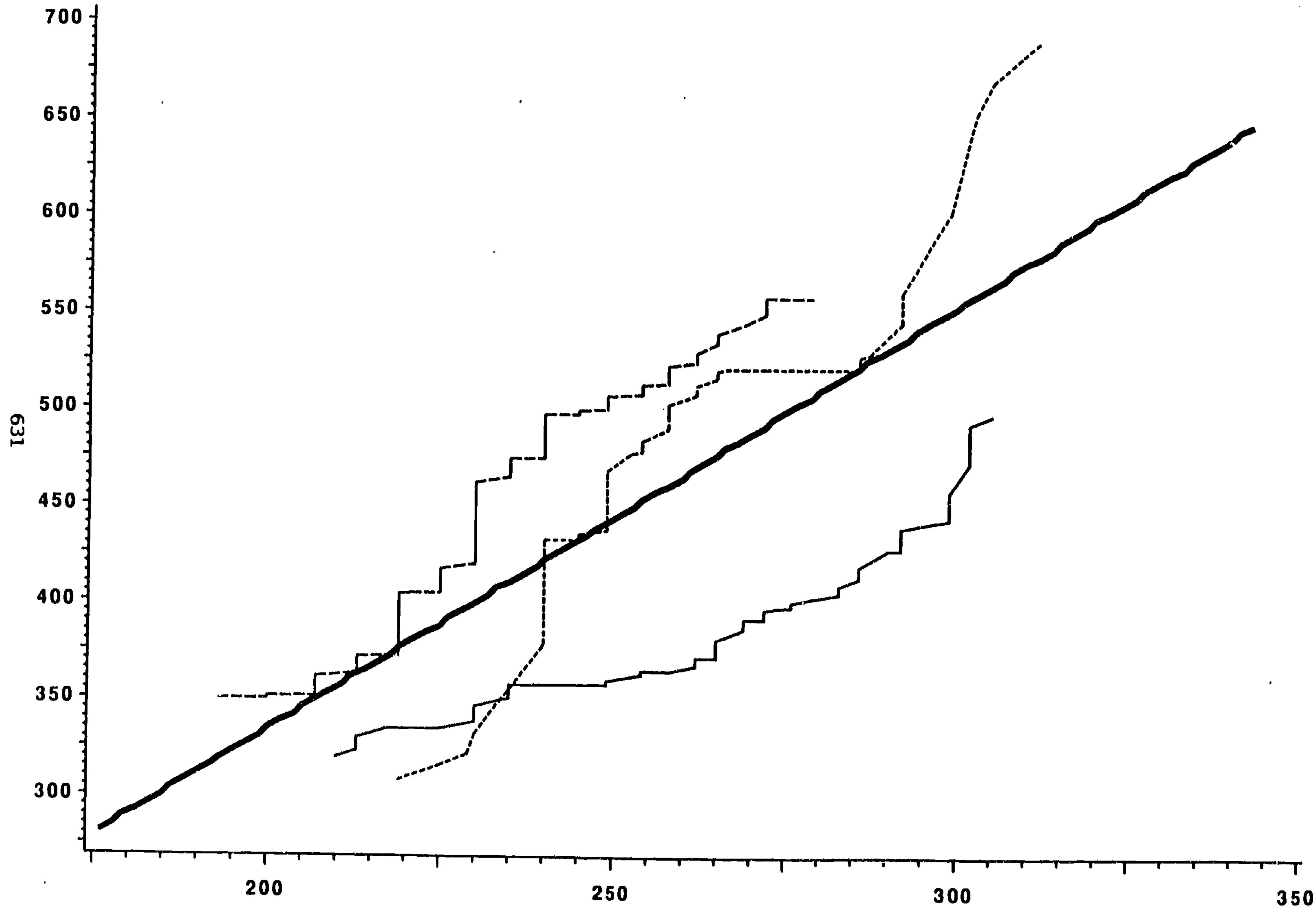


Figure 221

Math Grade 6 vs. Grade 1 (All Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in Districts D, E, & G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

B.	**Grade Span:**	**First Grade**	**to**	**Sixth Grade**
	Test Date:	**Spring**	**to**	**Spring**
	Language:	**English**	**to**	**English**
	Content:	**Language**	**to**	**Language**

As in mathematics, there is evidence that the average growth rate in English language skills of late-exit students seems to be comparable to that of this norming population (Figure 222). Closer inspection of the TAMP curve suggests that students with lower initial skills grew at a slightly faster rate than this norming population, while students with the highest initial skills grew at a slightly slower rate than this norming population. The latter may suggest the need to bolster the instructional program for the higher order language skills. Overall, contrary to what one would expect in the general population for disadvantaged students, these late-exit students appear to be holding their own relative to this norming population.

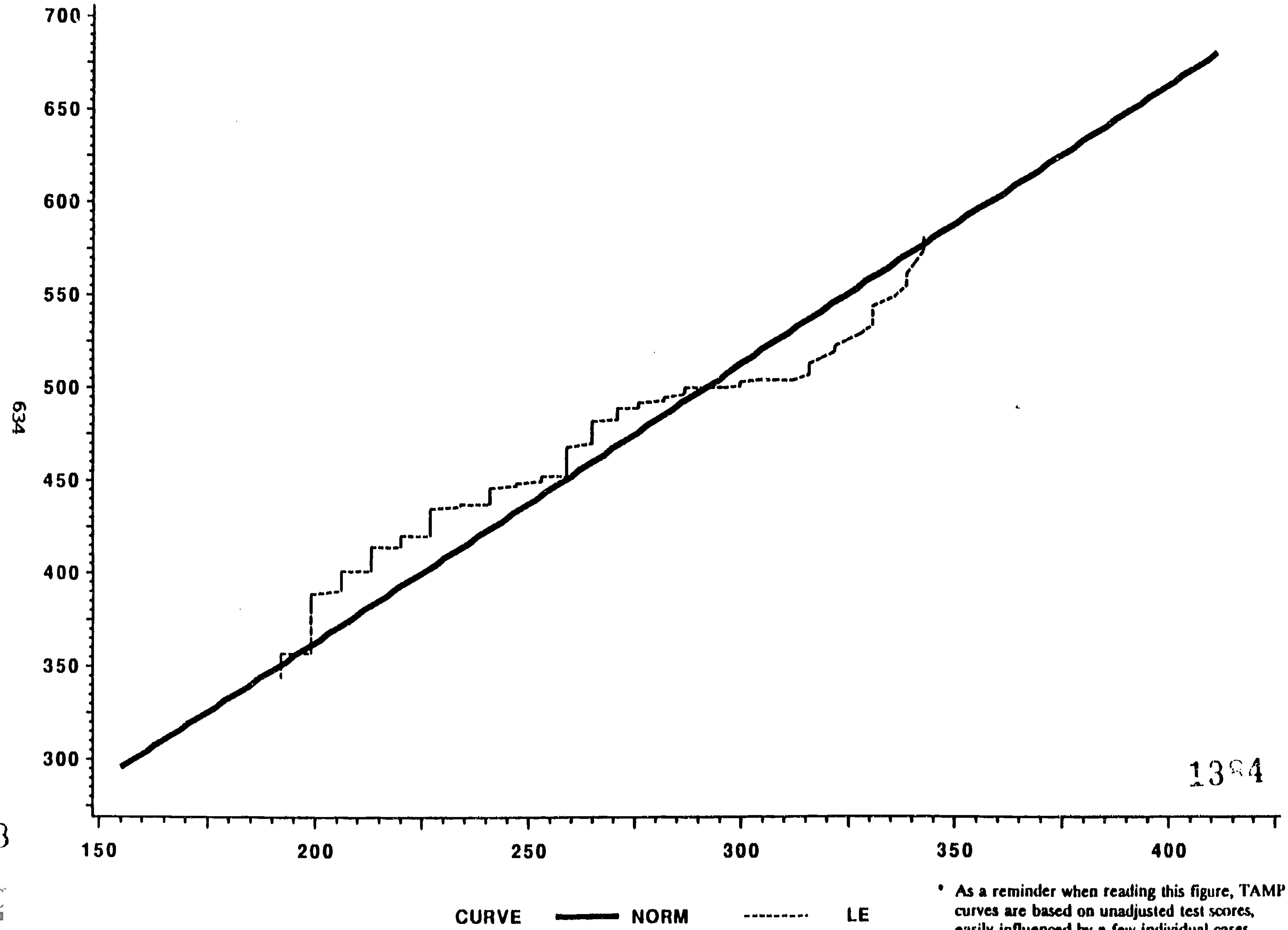
Once again, there is evidence that limited-English-proficient students who are provided with a strong primary language development program will acquire English language skills faster than students who are not. As before, site E students consistently seemed to grow at a faster rate from first grade to sixth grade relative to this norming population and the other two late-exit sites (see Figure 223). Site D students as a group appeared to approximate the growth rate of this norming population. As with mathematics, the growth pattern was uneven. Site D students with the lowest and highest initial English language skills seemed to learn at a slightly faster rate than this norming population, while site D students with average initial scores appear to have grown slightly slower relative to this norming population. Again this pattern may reflect the limited number of cases, resulting in a more unstable pattern.

Finally, site G students appear to have been losing ground. The data suggest that site G students with the lowest initial English language scores grew at about the same rate as this norming population. As the initial scores increased, their growth fell considerably below that of this norming population. This is of great concern for students at this school site. It suggests that the instructional program at this site only seems to be successfully meeting the needs of its lowest performing students. Students with stronger skills continue to lose their percentile rank relative to this norming population. Once again the growth patterns in English language skills mirror those found for mathematics. Students who were provided with more instruction in their primary language experienced the greatest growth in English language skills. Students who received about two-fifths of their instruction in the primary language were able as a group to keep pace with this norming population. Students with minimal instruction in their primary language and with abrupt changes from instruction in their primary language to instruction in English appear to experience a continual decline in their academic growth rate over time.

Figure 222

Language Grade 6 vs. Grade 1 (All Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program

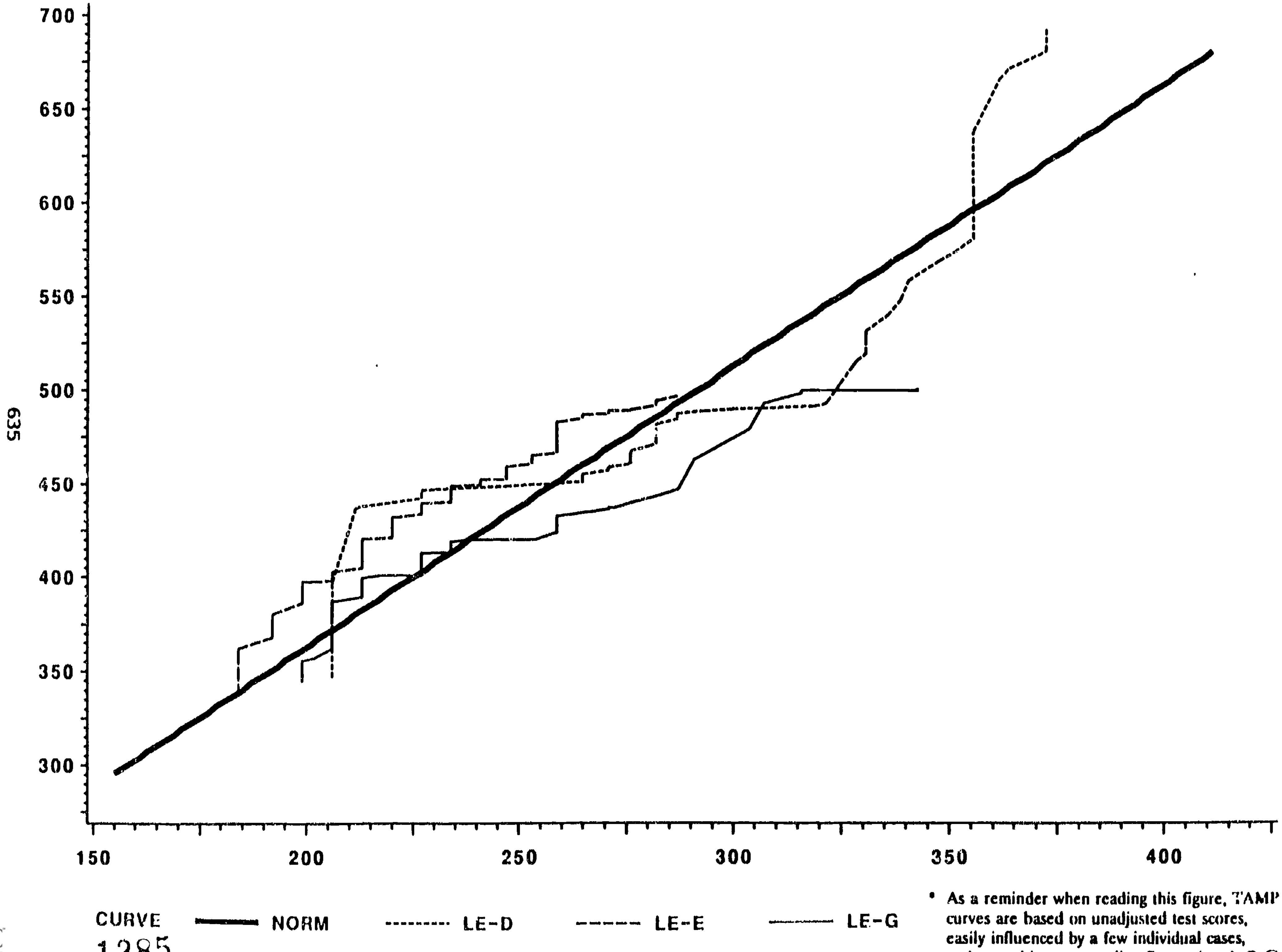


* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Figure 223

Language Grade 6 vs. Grade 1 (All Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in Districts D, E, & G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

C.	Grade Span:	First Grade	to	Sixth Grade
	Test Date:	Spring	to	Spring
	Language:	English	to	English
	Content:	Reading	to	Reading

As with mathematics and English language skills, late-exit students (collapsed across sites) appear to have improved their English reading skills at about the same rate as this norming population (see Figure 224). Students with low to average initial English reading skills seemed to grow at a slightly faster pace than this norming population. Students with the highest initial skills grew at a slightly slower pace than this norming population. There appear to be, however, sharp between-site differences in the rate at which English reading skills are developed (see Figure 225).

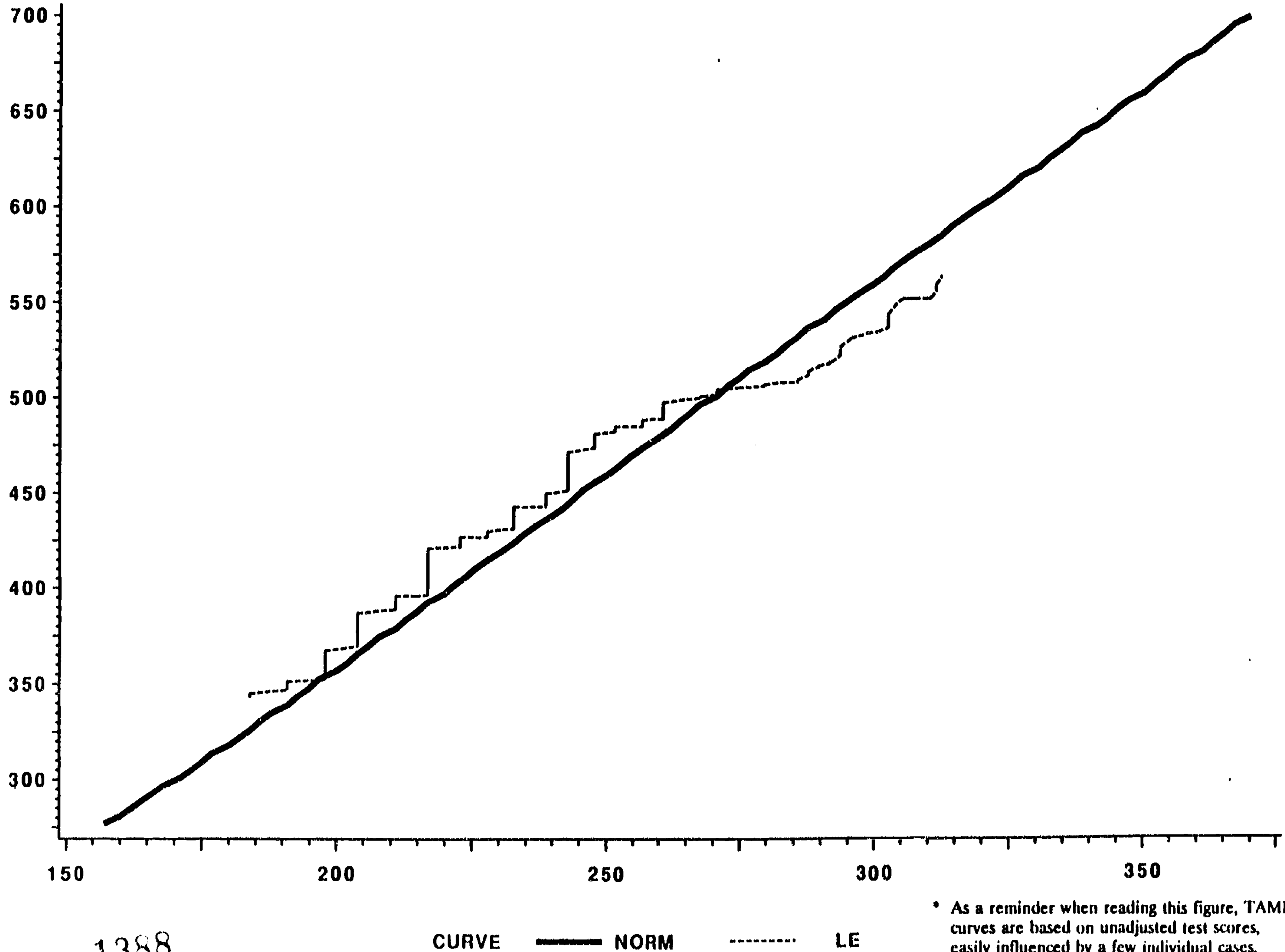
Once again, site E students at all initial skill levels seemed to grow at a much faster rate relative to this norming population. Site D students averaged a growth rate that appeared to be equal to that of this norming population. Among site D students, greater growth seemed to be realized by those students with the lowest initial skills. Site D students with average or higher skills seem to have been growing at the same rate as or slightly slower than this norming population. Site G students once again seemed to be consistently losing ground relative to this norming population. Of concern for this group is the trend observed for English language skills: the greater the skills, the slower the growth rate relative to this norming population.

ERIC

Figure 224

Reading Grade 6 vs. Grade 1 (All Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program



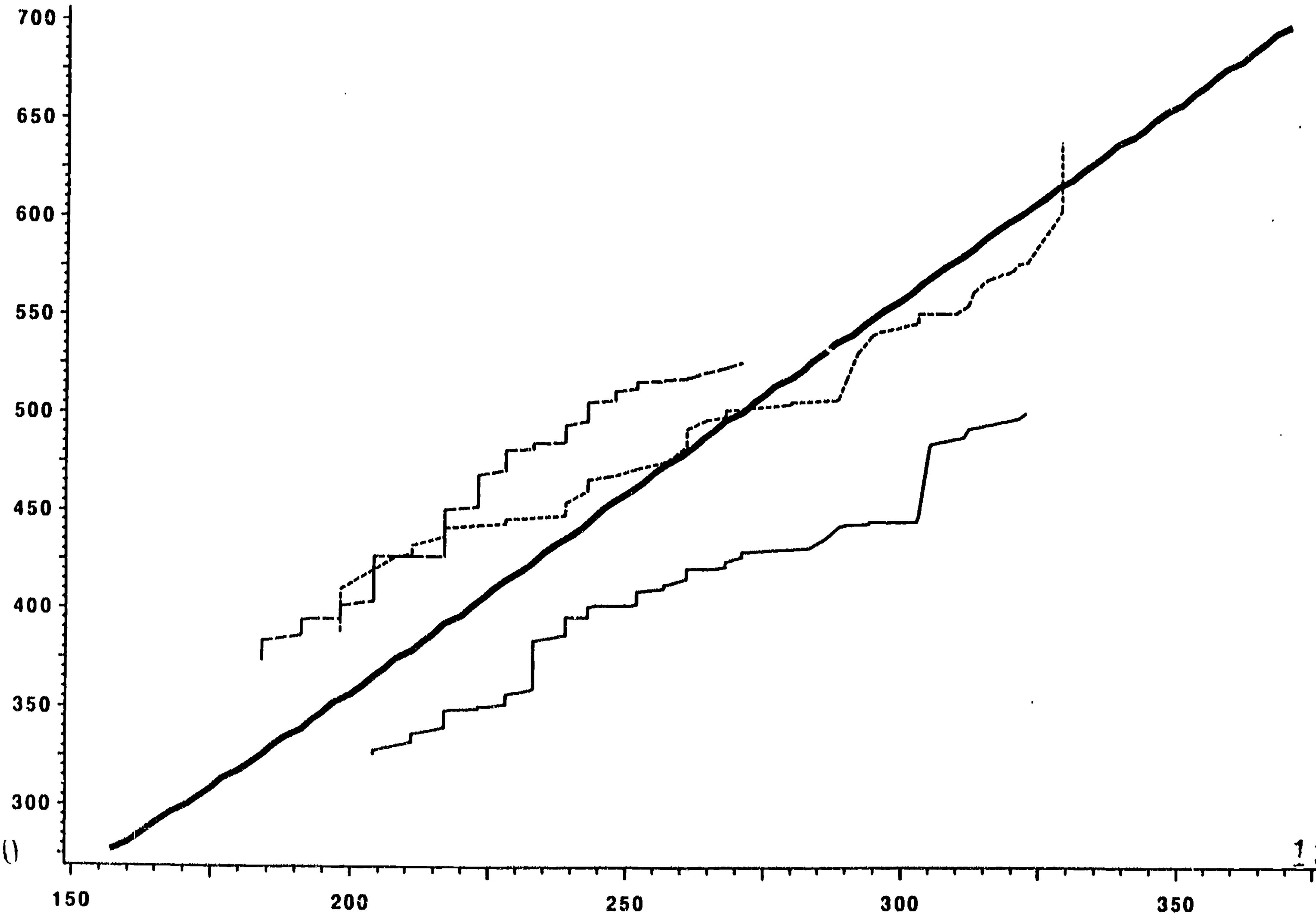
637

* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases, and are subject to sampling fluctuation.

Figure 225

Reading Grade 6 vs. Grade 1 (All Scores, Trim 5%)

TAMP Curves: Norming Population and Late-Exit Program in District E and District G



* As a reminder when reading this figure, TAMP curves are based on unadjusted test scores, easily influenced by a few individual cases.

Summary

To summarize, the first grade to sixth grade analyses provide an excellent summary and overview of the effect of the late-exit programs upon the learning of limited-English-proficient students in mathematics, English language, and English reading skills. The growth patterns are surprisingly consistent across content areas (see Table 263). This consistency in patterns lends support to the underlying hypothesis of the efficacy for providing limited-English-proficient students with a strong primary language development program. Over and over again, those students who began their schooling with substantial amounts of instruction in their primary language and were exposed to the gradual introduction of English for instruction realized the greatest growth in skills. Students who were exposed to a consistently even amount of instruction in their primary language were able to maintain their growth rate relative to that of this norming population. Students who began their schooling with about forty percent instruction in their primary language and were abruptly changed to almost exclusive instruction in English experienced a marked decrease in growth over time relative to this norming population. These students lost ground compared with this norming population, paralleling what is commonly observed for minority students in the general population. Thus, these TAMP analyses suggest that there is greater likelihood of moving the distribution of achievement scores for limited-English-proficient students closer to this norming population in mathematics, English language, and English reading skills if they are provided with a strong primary language development program. The data also strongly suggest that the worst instructional program for limited-English-proficient students is one with limited instruction in their primary language and an abrupt change in the language of instruction from their primary language to English. Such a program has a profoundly negative effect upon the ability of limited-English-proficient students to improve their skills in mathematics, English language, and English reading.

Table 263

Summary of TAMP Analyses Results:
Comparison of the Growth Rate for Each Late-Exit Site
Relative to the Growth Rate of This Norming Population

		Site D	Site E	Site G
1.	Kindergarten/Kindergarten (Fall/Spring)			
	A. Span-Mathematics/Span-Mathematics	=	=	*
	B. Span-Mathematics/Eng-Mathematics	+	-	*
	C. Span-Language/Span-Language	=	=	*
	D. Span-Language/Eng-Language	=	-	*
2.	Kindergarten/First Grade (Spring/Spring)			
	A. Span-Mathematics/Eng-Mathematics	+/=	+	+
	B. Eng-Mathematics/Eng-Mathematics	+/=	+	+
	C. Span-Language/Eng-Language	+	=	=
	D. Eng-Language/Eng-Language	+/=	+	+
	E. Span-Language/Eng-Reading	+/=	=	+
	F. Eng-Language/Eng-Reading	+	+	+
3.	Kindergarten/First Grade (Fall/Spring)			
	A. Span-Mathematics/Eng-Mathematics	+	+	*
	B. Span-Language/Eng-Language	+	=/-	*
	C. Span-Language/Eng-Reading	=/-	=/-	*
4.	First Grade/Third Grade (Spring/Spring)			
	A. Eng-Mathematics/Eng-Mathematics	+/=	=	-
	B. Eng-Language/Eng-Language	=	=/-	=/-
	C. Eng-Reading/Eng-Reading	=	+/=	-
5.	Third Grade/Sixth Grade (Spring/Spring)			
	A. Eng-Mathematics/Eng-Mathematics	=/-	=/-	-
	B. Eng-Language/Eng-Language	=	=	=
	C. Eng-Reading/Eng-Reading	-	-	=/-
6.	First Grade/Sixth Grade (Spring/Spring)			
	A. Eng-Mathematics/Eng-Mathematics	+/=	+	-
	B. Eng-Language/Eng-Language	=	+	=/-
	C. Eng-Reading/Eng-Reading	=	+	-

=: Growth rate equals/approximates that of this norming population

+: Growth rate is faster than that of this norming population

-: Growth rate is slower than that of this norming population

*: Test scores not available for fall kindergarten

Conclusion

The results of the TAMP analyses for grades one through three, with minor exceptions, are consistent with the findings for Chapters III through V. There appears to be no difference in the academic growth relative to the norming population between immersion strategy and early-exit students. Moreover, the form of this growth is similar to that found for late-exit students. When late-exit program sites are examined more closely, it appears that those sites that provided their students with the most primary language instruction consistently grew faster than this norming population. Those late-exit students who were provided with a consistent pattern of two-fifths of their instruction in their primary language seemed to grow at the same rate as this norming population. However, late-exit students who were provided with an abrupt transition into instruction exclusively in English seemed to lose ground relative to this norming population. This deceleration in growth is reminiscent of that observed for disadvantaged students in the general population.

VII. CONCLUSIONS AND IMPLICATIONS

Introduction

The purpose of this chapter is to summarize the findings presented in Volume II, relate them to the main study questions, and discuss their implications for instruction and policy. Specifically, this chapter addresses the following questions:

- What is the relative effectiveness of the immersion strategy and early-exit programs?
- What is the relative effectiveness among the three implementations of the late-exit instructional model?
- How does the academic growth of immersion strategy, early-exit, and late-exit students compare to the norming population used in this study?
- How does the academic growth of the three implementations of the late-exit program compare to this norming population?
- What recommendations can be made regarding the learning and instruction of language-minority students from these data?
- What do these data suggest for policy makers?

It should be noted that the programs participating in this study were carefully selected on the basis of their exhibiting those characteristics deemed critical to their respective instructional models. They represent the best implementations of their respective program models.

ERIC

Policy Context

The "Longitudinal Study of Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children" was prompted by the need for information on the kinds of instructional services that should be provided to language-minority children. The public policy question is, "What kinds of instructional services should be provided to limited-English-proficient students?"

Prior to this study there was insufficient empirical evidence to argue for or against the use of an adaptation of the Canadian immersion approach (i.e., structured English immersion) with minority language children or the use of a developmental primary language program (i.e., late-exit). Given this lack of information, the primary objective of this study was to assess the relative effectiveness of structured English immersion strategy, early-exit, and late-exit transitional bilingual education programs.

Effectiveness was defined in terms of student achievement. To address the study question, the analyses presented in this volume were designed to answer one and only one question: what differences in student achievement, as measured in English, were found among the three programs?

The data analyses were necessarily as complex as the study design. Reflecting the availability of immersion strategy, early-exit and late-exit programs in the field, a completely crossed design was not possible. No school district or school had all three programs. At best, schools and districts had both an immersion strategy and an early-exit program. Most schools offered only one of these two programs. Districts and schools that had a late-exit program provided only that instructional alternative. As a result, only immersion strategy and early-exit programs could be compared directly. In addition, the three late-exit program sites could be compared only with one another.

The grade spans covered by this study also differed by program model. A K-3 cohort and a 1-4 cohort of students were selected for both the immersion strategy and early-exit programs. These two cohorts were selected in anticipation that students would be mainstreamed after two or three years in the study, which would allow time to track the progress of mainstreamed students in the regular classroom. In contrast, a K-3 cohort and a 3-6 cohort of students were selected for the late-exit program. These two cohorts were selected to try and document the growth in skills of language-minority students who had received over five years of special instructional support.

As noted, the complexity of the study design precluded a direct analysis that concurrently assessed the relative effectiveness of all three programs. The analyses had to be effected in four separate steps. First, it was possible to assess the relative effectiveness of immersion strategy and early-exit programs where they existed within the same school. Second, a somewhat less direct analysis examined the relative effectiveness of the immersion strategy and early-exit programs in schools that had only one of these programs within districts that implemented both programs. Third, a separate analysis examined the effectiveness of three districts that had only a late-exit program in their schools. These three analyses were the most rigorous as they systematically considered the potential effect of non-program factors. Fourth, the academic growth of students in each of the three programs was compared to the growth of the national norms used in this study. This last set of analyses were primarily descriptive efforts directed towards illustrating the academic growt' f limited-English-proficient students in each program relative to this norming population. These analyses are limited in that they are based on unadjusted scores, i.e., potential effects of non-program factors were not considered.

Summary of Findings

Parts One and Two:
Relative Effectiveness of Immersion Strategy and Early-Exit Programs

The relative effectiveness of immersion strategy and early-exit programs was assessed in two steps; first in the analyses of two-program schools, and second in the analyses of one-program schools. The results of the two sets of analyses were fairly consistent. Because the two-program school analyses are more powerful, they are emphasized here.

- **Is achievement in mathematics comparable between structured English immersion strategy and early-exit transitional bilingual education students?**

Yes. Structured English immersion strategy and early-exit students do not differ in their mathematics skills or the rate at which they develop when tested in English. At the end of first, second, and third grade, students in the two programs have comparable skills in mathematics, according to both the two-program school analyses and the one-program school analyses.

- **Is achievement in English language skills as defined by the CTBS comparable between structured English immersion strategy and early-exit transitional bilingual education students?**

Yes. At the end of third grade immersion strategy and early-exit students demonstrated comparable English reading skills. However, differences in their growth rates were noted. While immersion strategy students had higher language skills than early-exit students at the end of first grade, by the end of third grade both groups were again comparable. This suggests that there is a temporary boost in language skills among immersion strategy students in first grade which decelerates thereafter, resulting in growth comparable to that of early-exit students from kindergarten through the end of third grade. Except for the temporary boost occurring in second grade rather than first grade, the one-program school analyses also found no difference in English language skills at the end of third grade.

- <u>**Is achievement in English reading skills as defined by the CTBS comparable between structured English immersion strategy and early-exit transitional bilingual education students?**</u>

Yes. While early-exit students had higher English reading skills than immersion strategy students at the end of first grade, by the end of third grade students in both programs read equally well. As in language, the immersion strategy students exhibited a boost in reading skills. However, this boost occurred in second grade rather than in first grade and appears to have been sufficient to bring them up to the level of the early-exit students. These results are evident in both the two-program school analyses and the one-program school analyses.

In sum, after four years in their respective instructional programs, immersion strategy and early-exit students demonstrate comparable skills in mathematics, language, and reading when tested in English.

Part Three:
<u>Effectiveness of Three Implementations of the Late-Exit Program</u>

The effectiveness of three implementations of the late-exit program was assessed following the general analytic steps used to compare immersion strategy and early-exit programs. As school districts that choose to implement a late-exit program do not provide either an immersion strategy or an early-exit program, it is not possible to compare these alternative instructional programs, nor is it possible to disentangle the effects of district and school from treatment effects. However, utilizing school and district level data it is possible to compare the three alternative late-exit programs with one another taking into account non-program factors that might have affected the results.

- **<u>Do students in the three late-exit sites have the same skills in mathematics?</u>**

No. At the end of third grade, students in all three implementations of the late-exit program had comparable skills in mathematics. However, by the end of sixth grade, the two late-exit sites that provided approximately forty percent or more of the instruction in the students' primary language (sites D and E) had significantly higher mathematics skills than students in the late-exit site who were abruptly transitioned into English instruction. The mathematics skills of students in the two late-exit sites with substantial primary language instruction did not differ.

- **<u>Do students in the three late-exit sites have the same skills in English language?</u>**

No. The late-exit site with the consistent level of primary language instruction (site D, at approximately 40%) and the highest level of English language scores at the end of first grade realized higher scores at the end of sixth grade than either of the other two late-exit sites. By the end of sixth grade, students in the remaining two late-exit sites (sites E and G) had virtually identical English language skills.

- **<u>Do students in the three late-exit sites have the same skills in English reading?</u>**

No. The late-exit program with the consistent level of primary language instruction (site D, at approximately 40%) and the highest level of reading scores at the end of first grade posted higher scores at the end of sixth grade than either of the other two late-exit sites (sites E and G).

In sum, there are differences between the three late-exit sites in achievement level for mathematics, English language, and English reading at the end of sixth grade. Students at the site with the highest skills in English language and reading in first grade (site D) also completed

sixth grade with the highest scores in these two areas. Students in the two remaining sites (the one with the most use of Spanish and the one with the most use of English) ended the sixth grade with the same skills in English language and reading. However, although all three late-exit sites had comparable mathematics skills in grade one, by the end of grade six, students in the two late-exit sites that used the most Spanish for instruction (sites D and E) posted higher growth than the site which abruptly transitioned into almost all English instruction (site G).

- **Do students in the three late-exit sites have the same growth in mathematics, English language, and English reading?**

Yes in English language and reading, and no in mathematics. While there may be differences between the three implementations of the late-exit program in English language and reading achievement as noted above, they do not differ in their rate of growth from first grade to sixth grade in these two areas. However, in the late-exit site wherein students were abruptly transitioned into English (site E), their rate of growth in mathematics was markedly lower (almost none) than the two late-exit sites providing forty percent or more of primary language instruction (sites D and E). What is important to note is that the growth for the students in the late-exit site wherein students were most at-risk and which provided the most primary language instruction (site E) was consistently greater than for the norming population. If this growth were sustained, in time their achievement would approach that of the norming population.

Students in all three late-exit sites show different growth rates between first and third grade than between third and sixth grade. The 1-6 analyses note that students in each of the three sites and in each content area realized greater growth in the early primary grades (K-3) than in the later primary grades (3-6).

There is wide variation in the effectiveness of the late-exit program between districts and schools within districts. Presumably this variation reflects the variation in implementation (i.e., proportion of English used for instruction) in the three late exit sites. The variation in results among the three late-exit sites and the apparent sensitivity of the late-exit model (as suggested by the within-program and between-program variation in results) strongly suggests the potential merit of effecting an operational analysis to try to identify those district, school, and classroom features that can be adjusted to maximize student learning.

There are differences in the growth curves between immersion strategy, early-exit, and late-exit students. The HLM analysis showed that the growth curve for immersion strategy and early-exit students was negative, indicating a deceleration in their rate of growth from first grade to third grade. In contrast, the growth curve for late-exit students was positive from grade one to grade three, suggesting continued growth over this grade span.

Part Four:

Comparison of the Academic Growth of Immersion Strategy, Early-Exit, and Late-Exit Students With the Norming Population Used in This Study

The TAMP analyses were completed to allow a comparison of the growth of each of the three instructional programs to the norming population used for this study. As the scope of work required that the test data for this study had to be comparable to that of other federal studies (e.g., Sustaining Effects Study), an older version of the CTBS was used. The norming population used in the TAMP analyses is the norming population for that version of the CTBS. If other norming populations were to be selected, while the form of the growth patterns would not change, their relative position to the norming population curve might. In addition, the growth of the norming population is based on a cross-sectional sample rather than a longitudinal sample, which causes a number of difficulties. Other limitations to the TAMP analyses are that they are based on unadjusted test scores, i.e., non-program factors are not considered. With these caveats in mind, the TAMP analyses provide a good description of the patterns in academic growth for limited-English-proficient students.

o **Is there a difference in the rate at which immersion strategy, early-exit and late-exit students increase their mathematics skills relative to the norming population used in this study from kindergarten to third grade?**

Yes and no. It appears that from kindergarten to first grade, students in all three programs increased their skills in mathematics as fast as or faster than the norming population used in this study. From first to third grade, however, it seems that students in each program experienced a growth in mathematics that was slower than this norming population.

o **Is there a difference in the rate at which immersion strategy, early-exit and late-exit students increase their English language skills relative to the norming population used in this study from kindergarten to third grade?**

No. It seems that from spring kindergarten to spring first grade and from spring first grade to spring third grade, students in all three programs increased their skills in English language as fast as or faster than this norming population.

o **Is there a difference in the rate at which immersion strategy, early-exit and late-exit students increase their English reading skills relative to the norming population used in this study from kindergarten to third grade?**

No. As with English language skills, it appears that students in all three programs realized a growth in reading skills that was as fast as or faster than this norming population.

o **Is there a difference in the rate at which students in each of the late-exit sites increased their skills in mathematics relative to the norming population used in this study from kindergarten to sixth grade?**

Yes. The TAMP curves suggest that students in site E, who were provided with substantial instruction in their primary language and a slow phasing in of English instruction over time, consistently realized the

ERIC

greatest growth in mathematics skills, faster than this norming population. Students in site D, who were exposed to a consistent proportion of instruction in their primary language (approximately 40%), realized growth in mathematics that was equal to this norming population. Noteworthy is that after covariates were considered, there was no difference in the achievement of students in sites D and E, although students in site E had more stress in their environment and fewer resources than site D students.

In contrast, it appears that students in site G who received about 40% of their instruction in their primary language in kindergarten and first grade, but were then abruptly moved into almost exclusive instruction in English (comparable to that provided in early-exit and immersion strategy programs), experienced a marked decrease in growth in mathematics skills over time relative to this norming population. It seems that these students lost ground relative to this norming population, paralleling what is commonly observed for disadvantaged students in the general population (Bureau of the Census, 1985; Burton and Jones, 1982; Escutia and Prieto, 1987; Levin, 1986; National Center for Education Statistics, 1989; National Commission on Secondary Education for Hispanics, 1984; Orum, 1986; Rumberger, 1983; United States General Accounting Office, 1987; Vargas, 1988).

- **Is there a difference in the rate at which students in each of the late-exit sites increased their English language skills relative to the norming population from kindergarten to sixth grade?**

Yes. As in mathematics, by the end of sixth grade it appears that the late-exit students in site E, who started below average and had had the most opportunity to develop their primary language, increased their English language skills faster than the norming population used in this study. If sustained over time, this would suggest that these students are gaining on this norming population and would eventually approximate the average English language achievement level of this norming population. Students in site D, who received less, although considerable, primary language instruction, appeared to increase their English language skills at the same rate as this norming population. It seems that these students are keeping their position relative to this norming population. Site G students, who were transitioned quickly into English in second grade, appeared to exhibit a decrease in their rate of growth in English language skills, suggesting that these students, regardless of their skill levels in third grade, were losing ground relative to this norming population.

- **Is there a difference in the rate at which students in each of the late-exit sites increased their English reading skills relative to the norming population from kindergarten to sixth grade?**

Yes. As in mathematics and English language, it seems that those students in site E, who received the strongest opportunity to develop their primary language skills, realized a growth in their English reading skills that was greater than that of the norming population used in this study. If sustained, in time these students would be expected to catch up and approximate the average achievement level of this norming population. It appears that once again site D students, who were provided with two-fifths of their instruction in their primary language through grade five, seemed to increase their English reading skills at the same rate as this norming population. That is, they kept up with this norming population. Site G students, who were abruptly transitioned into English instruction, appeared to experience a slower increase in English reading skills from third grade to sixth grade relative to this norming population; they seemed to be losing ground.

In sum, a consistent pattern seems to be emerging in the TAMP figures. It appears that students who were provided with a substantial and consistent primary language development program learned mathematics, English language, and English reading skills as fast as or faster than the norming population used in this study. As their growth in these academic skills is atypical of disadvantaged youth, it provides support for the efficacy of primary language development in facilitating the acquisition of English language skills. The TAMP curves using unadjusted scores suggest that limited-English-proficient students who are abruptly transitioned into an English-only instructional program appear to lose ground (in terms of decelerating rate of growth) relative to this norming population in all three content areas, a pattern which is consistent with the growth of disadvantaged students in the general population (Bureau of the Census, 1985; Burton and Jones, 1982; Escutia and Prieto, 1987; Levin, 1986; National Center for Education Statistics, 1989; National Commission on Secondary Education for Hispanics, 1984; Orum, 1986; Rumberger, 1983; United States General Accounting Office, 1987; Vargas, 1988). While the HLM analysis confirms the lower growth in mathematics for students who are transitioned abruptly, it notes that the growth rates in language and reading are the same as those of students in the two late-exit sites wherein substantial amounts of primary language instruction are provided. Nonetheless, if the pattern of observed growth rates are sustained over time, students with substantial amounts of primary language instruction would be expected ultimately to outperform those students who are transitioned quickly into English instruction in English language and reading skills.

These findings suggest that providing LEP students with substantial amounts of instruction in their primary language does not impede their acquisition of English language skills, but that it is as effective as being provided with large amounts of English. Of equal importance is the finding that students who are provided with substantial amounts of primary language instruction are also able to learn and improve their skills in other content areas as fast as or faster than the norming population, in contrast to students who are transitioned quickly into English-only instruction.

Discussion

There are four major findings drawn from this study:

1. There is no difference in the level of achievement or rate of growth in achievement between students in an immersion strategy program and an early-exit program after four years in their respective programs (i.e., end of third grade).

2. Limited-English-proficient students can be provided with substantial amounts of primary language instruction without impeding their acquisition of English language and reading skills.

3. Limited-English-proficient students who are provided with substantial instruction in their primary language ($\geq$ 40%) successfully continue to increase their achievement in content areas such as mathematics, while they are acquiring their skills in English; in contrast, students who are quickly transitioned into English-only instruction tend to grow slower than the norming population.

4. Students in all three late-exit instructional programs appear to exhibit greater growth from spring of first grade to spring of third grade than from spring of third grade to spring of sixth grade. This deceleration in growth mirrors that of the norming population.

All of these findings are consistent with theory and available research. While no data are available documenting the efficacy of an immersion strategy program with limited-English-proficient students, limited data are available describing the academic growth of language-minority students in a primary language development program, i.e., late-exit. Recall from Chapter I that for the purposes of this study, the late-exit model, rather than the early-exit model, is considered the optimum example of a primary language development program reflecting the greater amount and duration of primary language instruction.

Two lines of research are relevant. The first line is based upon studies that focus on the academic growth of limited-English-proficient students in primary language development programs. The second group of studies are those that examine the relationship of student age and the acquisition of second language skills.

Studies of the Academic Growth of Limited-English-Proficient Students

Cumulative effects of primary language instruction. The first of these studies examines the bilingual program provided to Navajo students in Rock Point in the United States. The primary language of instruction for students in kindergarten and first grade was Navajo. Formal instruction in English reading did not begin until second grade, after students had learned to read in Navajo. Moreover, from second grade to sixth grade both Navajo and English were used for formal instruction. When the achievement of these students was compared to that of comparable Navajo students who had not received instruction in Navajo, it was concluded that:

> "Data presented suggest that the effects of continuous bilingual instruction may be cumulative, that while Navajo students who have recently (in second grade) added reading in English to reading in Navajo may not do better on standardized achievement tests than Navajo students who began reading in English, they do achieve better test scores each year thereafter. Nor does the difference seem to remain the same. The students who learned to read in Navajo and who continue to learn through Navajo and English appear to obtain scores progressively higher in English than those who did not. In effect, their rate of growth helps them to achieve progressively closer to the "national norms" in each grade third through sixth, instead of maintaining a "continuously retarded" level of achievement" (Rosier and Farella, 1976).

The findings of the study discussed in this report replicate these observations. The HLM analyses revealed no difference in achievement between the immersion strategy or early-exit students, nor between the students in the three implementations of the late-exit model at the end of third grade. Moreover, the TAMP analyses find little difference in the growth of academic skills among immersion strategy, early-exit, and late-exit students from first grade through third grade. However, from third grade through sixth grade noticeable differences in the rate of growth are evident between the three late-exit models, with those students who received substantial amounts of primary language instruction exhibiting faster growth than students who were quickly transitioned into English-only instruction.

Operationally, one of the three late-exit programs was not as faithful to the late-exit program model. This site abruptly transitions children from instruction in the primary language into almost exclusive use of English. This site allows us the opportunity to examine the potential effects of quickly transitioning a limited-English-proficient student from a primary language development program into one wherein English is used almost exclusively for instruction.

The 1-6 TAMP analysis are consistent with the observations of the Rock Point project. Those students who had the most opportunity to develop their primary language skills (i.e., in sites D and E, with more primary language instruction) grew at a faster rate in mathematics than the norming population. If this growth rate were sustained over time, one would expect the average mathematics performance of these students to approximate that of the norming population. The importance of this result is even more impressive given the school district where one of the programs was implemented. Of all study sites, site E had families, schools, and community with by far the greatest needs and the fewest resources to address these needs. Yet these students and those from site D realized the greatest growth in mathematics over time as compared to the norming population and the late-exit site where students were abruptly transitioned into English instruction (site G).

This growth in mathematics by students receiving substantial amounts of primary language instruction is of particular interest in that mathematics is not as language-dependent as are English language and reading. The growth in mathematics by students with substantial primary language instruction supports the hypothesis that LEP students can

ERIC

continue to grow in the content areas while they are developing their English language skills. In contrast, students who were abruptly transitioned into an English instructional environment did not grow as fast, but appeared to experience a decrease in their rate of growth in mathematics relative to the norming population.

The potential effect of being provided with less primary language instruction is further illustrated by examining the growth of students in the remaining two late-exit sites (sites D and G). Students who were provided with a constant forty percent of their instruction from kindergarten through fifth grade (site D) grew at the same rate as the norming population. This suggests that students in such a program would be able to maintain their position relative to the norming population. In short, they would not demonstrate a decrease in their rate of growth as is so often the case with minority children. Once again, this finding reinforces and is consistent with the Rock Point conclusions.

That students who were transitioned quickly into English instruction (site G) lost ground in mathematics relative to the norming population and to the other two late-exit sites serves once again to underscore the value of primary language instruction.

That the rate of growth in English language and reading was the same across the three sites demonstrates that achievement of LEP students in these two subject areas is not impeded by being provided with substantial instruction in their primary language.

Study findings are also consistent with results reported outside of the United States. Skutnabb-Kangas (1981) reported that Finnish students who completed their initial schooling in Finland realized greater success in learning Swedish after two years of study than Finnish children who were instructed in Swedish since the first grade.

Higher proficiency in L1 facilitates acquisition of L2. In an innovative study, Swain, Lapkin, Rowen, and Hart (in press) examined the question of whether literacy in one's first language facilitated the acquisition of a third language. This study is pertinent as it examines the importance of the development of strong primary language skills in the learning of another language.

Swain et al. (in press) compared the French language proficiency (i.e., reading, writing, comprehension, and speaking) of native English speakers with that of language-minority students (e.g., Spanish, Tagalog, Vietnamese, etc.) at the end of grade eight. Language-minority students were subdivided into two groups: those that could speak, read, and write in their home language and those that could not. Students in each of the three groups had received all of their instruction in English, with approximately twenty minutes per day in French as a second language class through grade four. Beginning in grade five, each of the three student groups received half of their instruction in English and half in French through grade eight.

Results show that literacy in one's first language, "regardless of whether learners are currently making use of those literacy skills, has a strong positive impact on the learning of a third language" (Swain et al., in press, p. 116). "Our results suggest that the effect is related to literacy knowledge (whether currently used or not) rather than oral proficiency in the HL" (Swain et al., p. 117). To try and clarify these findings, Swain et al. examined the test scores to determine the extent to which they might simply be a reflection of a general high level of proficiency in the home language or a result of being literate in the home language. To this end, the relationships to test scores of frequency of use and literate and nonliterate background were examined. They concluded that literacy in the home language facilitates the learning of a third language and that it is independent from that of a general proficiency in the home language. "What is interdependent is knowledge and process" (Swain et al., p. 120). These results were not affected by various indicators of socioeconomic status such as parent income, education, or occupation.

Results of this study are consistent with the findings of Swain and her colleagues. The HLM analyses found that those students from the late-exit site with the highest initial skills in English language and reading also realized the highest skills at the end of the sixth grade. From the TAMP analyses, those students in the one late-exit site who had received the most instruction in and opportunity to develop their primary language skills experienced the greatest growth, faster than the norming population, and faster than students in the other two late-exit sites.

When L2 should be learned. Study findings are not entirely consistent with Canadian research regardi when L2 should be learned. Genesee (1985) compared the proficiency in French at the end of sixth grade of two groups of students who had English as their primary language. In one group, students had received all of their instruction in French since the first grade (early immersion). The second group received all of their instruction in English through grade three; in grade four, half of their instruction was in English and half in French (delayed immersion). While the two groups of students began school with comparable low levels of French, early immersion students scored higher in French skills at third grade than delayed immersion students. In contrast, in this study the HLM analyses found no difference in the achievement level in mathematics, English language or reading between LEP students who had been exposed to an early immersion program (i.e., immersion strategy) or to a delayed immersion program at the end of third grade. Presumably, the difference between the Genesee results and those of this study is that students in this study came from homes that were not as advantaged educationally and economically as those in the Genesee study, and therefore had primary language skills that were not as well developed as those in the Genesee study; as a result, these students were not able to learn the second language as fast as students in the Genesee study.

The two studies are consistent when the growth of LEP students is considered from grades one through six. Surprisingly, in the Genesee study, when tested at the end of sixth grade, students who did not begin instruction in French until grade four were just as successful in

ERIC

acquiring skills in French as those students who had been instructed in French since their first day of school. Genesee attributed the effectiveness of the delayed immersion program to the cognitive and linguistic maturity of the older students. That is, developing these skills in their first language facilitates (requires less time) acquiring these skills in a second language. This is consistent with the TAMP analyses, which note the faster growth rate in English language and reading as well as in mathematics by students receiving substantial amounts of primary language instruction than students who were quickly transitioned into English-only instruction.

Age-Related Factors and the Acquisition of Second Language Skills

Collier's (1988) literature review specifically addressed the issue of age and the acquisition of second language skills. She gathered information on the effect of age on basic oral skills, oral and writing skills for school, and language skills in the development of subject matter content area skills (e.g., social studies, science, etc.). Studies which focused on the short-term and long-term effects of exposure to a second language suggest that children over the age of seven acquire basic oral skills much more quickly in a second language than younger children; however, over time the younger learners eventually catch up and surpass the language proficiency of the older learners. When one considers the more demanding use of language in the classroom, students over the age of seven "are faster, more efficient acquirers of school language than younger students" (Collier, 1988, p. 3). Collier proposes that, as children are still in the process of developing and consolidating their primary language skills through twelve years of age, trying to get them to acquire a second language is very difficult and requires more time than it does for older students who have already developed their primary language skills and can simply transfer their skills into the second language. Moreover, theorists note and other researchers demonstrated that older students are more cognitively mature and are able to bring to the task of learning a second language a greater repertoire of strategies for acquiring a new language than younger learners.

In sum, the rate at which younger students acquire a second language cannot be rushed. Results of this study are consistent with Collier's (1988) findings. The HLM analyses did not reveal any differences in the rate of growth nor in the achievement level in English language and reading between immersion strategy and early-exit students. This suggests that "immersing" LEP students in English-only instruction does not hasten the rate at which they can develop their English skills. This is reinforced by the TAMP analyses which show little difference in the growth in mathematics, English language, or reading skills between immersion strategy, early-exit, or late-exit students relative to the norming population. The TAMP analyses of the three late-exit models from grade one through grade six also suggest that providing LEP students with more English (through an abrupt transition into English-only instruction) may in fact serve to slow the rate at which LEP students can acquire skills not only in English, but in other content areas as well.

Child Development Theory

As noted earlier, the findings from this study are also consistent with child development theory. Twyford's (1988) discussion of age-related factors that might affect second language learning is quite helpful. He appropriately reminds us of Piaget's conceptualization of cognitive development. Succinctly, Piaget's observations of children's speech led him to posit that children shift from "egocentric" to "socialized" speech at about six or seven years of age. This change is one from where children move from "monologues" (talking without regard to a listener) to actual verbal exchanges with others. This shift reflects a change in the child from the pre-operational to the concrete operational stage of cognitive development. This conceptualization of language suggests that children over the age of seven would respond more successfully to formal instruction in a second language than children who are younger than eight years. Assuming that kindergarten children are five years old, third graders would be eight years old. Thus, while one would not expect students below third grade to respond successfully to formal instruction in a second language, students at and beyond fourth grade would be expected to show some success. It would follow, therefore, that any attempts to facilitate the acquisition of a second language at the K-3 level would not be very successful, regardless of the method. This is consistent with the finding from the HLM analyses that there are no differences in the achievement levels of mathematics, language, and reading skills when tested in English between immersion strategy and early-exit students at the end of third grade. Moreover, this conceptualization of language development also suggests that there would be no difference in the kindergarten through third grade growth rates in these skills between immersion strategy, early-exit, and late-exit students. The TAMP results are consistent with these projections.

Piaget's view of language development would posit that an increase in English language skills should begin to occur around the fourth grade and continue to increase thereafter. When the growth rates of students in the three late-exit sites are examined, this increase is noted for two of the sites, those providing limited-English-proficient students with substantial instruction in their primary language.

In sum, the theoretical formulations provided by Piaget offer a conceptual framework for explaining the consistent patterns found not only in this study, but in studies of the success of limited-English-proficient students in acquiring a second language.

Implications

Implications for Pedagogy

- **Limited-English-proficient students in all three instructional programs improved their skills in mathematics, English language, and reading as fast as or faster than students in the general population. Providing substantial instruction in the child's primary language does not impede the learning of English language or reading skills. On the other hand, providing a limited-English-proficient student with English-only instruction through grade three, as was done in the structured English immersion strategy program, is as effective as an early-exit program in helping limited-English-proficient students acquire mathematics, English language, and reading skills.**

- **When immersion strategy programs are staffed with teachers who have specialized training in bilingual education (bilingual teaching credential) and English as a second language methodology (language development specialist or ESL teaching credential), they are viable alternatives to early-exit transitional bilingual education programs through grade three.**

The achievement in mathematics, English language, and English reading skills for LEP students in immersion strategy programs demonstrate a temporary boost in the growth of English language skills (in first or second grade) and possibly also in reading. Thus the immersion strategy students have achieve at least as high as early-exit students by the end of third grade.

- **Early-exit transitional bilingual programs as implemented in this study are effective in meeting the needs of language-minority students in grades kindergarten through three.**

The academic growth of early-exit LEP students in mathematics, English language arts and reading suggests continuing growth from kindergarten through grade three. The growth exhibited by LEP students in the early-exit program is comparable to that demonstrated by the CTBS norming population used in this study, the majority of whom are native English speakers. Moreover, the increased achievement by early-exit LEP students is better than one would have expected, given their low entry level skills.

- **Late-exit transitional bilingual education programs, particularly those that are most faithful to the late-exit instructional model, are viable alternative instructional education programs.**

Late-exit students exhibit growth rates higher than the national norms in mathematics, English language, and English reading from first grade through sixth grade, with the growth rates in grades one to three being higher than in grades three to six.

From another perspective, while the growth curves for immersion strategy and early-exit students show growth from first grade to third grade in mathematics, English language, and reading skills, they also show a slowing down in the rate of growth in each of these content areas as grade level increases. This deceleration in growth is similar to that observed for students in the general population. In contrast, the growth curves for late-exit students from first grade to third grade and from third grade to sixth grade suggest not only continued growth in these areas, but continued acceleration in the rate of growth, i.e., they are gaining on students in the general population.

- **There is limited support for the hypothesis that development of a language-minority student's primary language facilitates the acquisition of English language skills and of other content areas. Moreover, providing students with substantial instruction in their primary language does not impede their progress in mathematics.**

It is hypothesized that the development of these skills will not be as strong in the early primary grades (i.e., kindergarten through third grade) as it will be in the later middle school grades (i.e., fourth grade through sixth grade). Consistent with the predictions of student achievement for late-exit students, it seems that these students did not exhibit as much growth in English language or reading skills relative to the norming population used in this study as did immersion strategy or early-exit students in kindergarten through third grade. However, it appears that the late-exit growth was comparable to that of this norming population for mathematics.

Although students in the late-exit program who started out lower than average grew slower than the norming population in kindergarten to third grade, they grew faster than the national norms in grade three to grade six. This is consistent with the underlying hypothesis of the late-exit model. This suggests that, if this slope were maintained, over time the late-exit students would approximate the average achievement level of the norming population. These results tend to support the notion that development of primary language skills will facilitate the acquisition of English and other content areas.

The efficacy of late-exit programs is given additional impetus when the growth curves of the students in the two late-exit sites that most closely resemble the late-exit instructional model are compared to the growth curve of those students in the late-exit site that provided an abrupt transition into an English instructional program. This transition is not unlike that experienced by most limited English proficient students wherein they are quickly shifted from using their primary language for instruction into almost the exclusive use of English. In light of this similarity, one might hypothesize that the growth rates at late-exit site G might reflect the growth curves of students at immersion strategy and early-exit programs, had comparable data been available for these two programs. The academic growth of students at site G appears to be consistently lower in mathematics skills than the norming population used in this study. In marked contrast, it appears that those late-exit students who received the most primary language instruction grew faster than this norming population. It seems that those students who consistently received about two-fifths of their instruction in their primary language grew as fast as this norming population. These findings for the two "true" late-exit programs seem to indicate once again that if this growth rate were sustained, over time these late-exit students would either catch up to and approximate the average achievement of this norming population, or at least keep their position relative to this norming population. Neither group would lose ground, as is typical for disadvantaged students. This suggests the efficacy of the late-exit instructional model as implemented by these two sites.

- **There is limited support for the hypothesis that in the early primary grades (i.e., kindergarten through third grade), the greater the proportion of instruction provided in English (as exhibited in immersion strategy), the greater the growth in English language arts skills among LEP students.**

However, the ambiguity of the results in English reading suggests that the benefits of greater exposure to instruction in English in kindergarten through grade three may be limited to the form of language and not the deeper conceptual meaning of language. This again is consistent with second language research that states that "students reach the fiftieth normal curve equivalent on a standardized test [of English language grammar] two to four years earlier than they reach the fiftieth normal curve equivalent on the reading test" (Collier, 1990, p. 521).

Implications for Public Policy

- **If the concern is for the achievement of LEP students in the short run, i.e., through third grade, instructional services can be provided through either immersion strategy or early-exit programs.**

There is no difference in the achievement between immersion strategy and early-exit students after four years of instruction in their respective programs. However, how the two programs compare after third grade is unknown.

- **Current federal efforts to support primary language (i.e., developmental bilingual) programs that provide substantial amounts of primary language instruction as implemented in some of the late-exit programs in this study (sites D and E) is justified.**

Exposure to primary language development programs as reflected in some of the late-exit programs in this study does not inhibit the growth of mathematics, English language, and English reading skills. On the contrary, there is some suggestion that the academic growth of these students is faster than that of the norming population used in this study. Further analyses at the operational level are needed to identify those classroom, school, and district characteristics that are critical to student success.

- **There is evidence that suggests that when limited-English-proficient students receive most of their instruction in their home language, they should not be abruptly transferred into a program that uses only English.**

REFERENCES

Braun, H.I. (198ſ . A new approach to avoiding problems of scale in interpreting trends in mental measurement data. Journal of Educational Measurement, 25(3), 171-191.

Bryk, A.S. & Raudenbush, S.W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. American Journal of Education, 97(1), 65-108.

Bryk, A.S., Raudenbush, S.W., Seltzer, M., & Congdon, R.T., Jr. (1988). An introduction to HLM: Computer program and users' guide. Unpublished.

Bureau of the Census, United States Department of Commerce (1985). Education in the United States: 1940-1983. Washington, DC: United States Government Printing Office.

Burton, N. & Jones, L. (1982). Recent trends in achievement levels of Black and White youth. Educational Researcher, 11, 10-14.

Cannell, J. (1988). Nationally normed elementary achievement testing in America's public schools: How all 50 are above the national average. Educational Measurement: Issues and Practice, 7(2), 5-9.

Collier, V. (1988). The effect of age on acquisition of a second language for school. New Focus, 2, 1-7.

Collier, V. (1990). How long? A synthesis of research on academic achievement in a second language. TESOL Quarterly, 23(3), 509-531.

Genesee, F. (1985). Second language learning through immersion: A review of U.S. programs. Review of Educational Research, 55(4), 541-561.

Levin, H. (1986). Educational Reform for Disadvantaged Students: An Emerging Crisis. West Haven, CT: National Education Association of the United States.

Linn, R., Graue, E., & Sanders, N. (1990). Comparing state and district test results to national norms: The validity of claims that "everyone is above average." Educational Measurement: Issues and Practice, 9(3), 5-14.

Lord, F.M. (1982). The standard error of equipercentile equating. Journal of Educational Statistics, 7, 165-174.

McGraw-Hill, Inc. (1974). Comprehensive Tests of Basic Skills, Form S, Technical Bulletin No. 1. Monterey, CA: CTB/McGraw-Hill.

National Center for Education Statistics (1989). The Condition of Education. Washington, DC: United States Government Printing Office.

National Commission on Secondary Education for Hispanics (1984). Make Something Happen: Hispanics and Urban School Reform, Volume II. New York, NY: Hispanic Policy Development Project.

Orum, L. (1986). The Education of Hispanics. Washington, DC: National Council of La Raza.

Ramirez, J.D., Schinke-Llano, L., & Bloom, G. (1984). The characteristics and effectiveness of transitional bilingual education programs. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J.D., Wolfson, R., & Morales, F. (1985). Description of immersion strategy programs in the United States. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J.D., Wolfson, R., Tallmadge, G.K., & Merino, B. (1984). Study design of the longitudinal study of immersion programs for language-minority children. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J.D., Yuen, S.D., & Ramey, D.R. (1987). Second year report: Longitudinal study of immersion, early-exit and late-exit transitional bilingual education programs for language-minority children. Mountain View, CA: SRA Technologies, Inc.

Ramirez, J.D., Yuen, S.D., & Ramey, D.R. (1988). Third year report: Longitudinal study of immersion, early-exit and late-exit transitional bilingual education programs for language-minority children. San Mateo, CA: Aguirre International.

Rogosa, D.R. (1979). Causal models in longitudinal research: Rationale, formulation and interpretation. In J.R. Nesselroade & P.B. Baltes (Eds.), Longitudinal research in the study of behavior and development (pp. 263-302). New York: Academic Press.

Rogosa, D.R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 92, 726-748.

Rogosa, D.R., & Willett, J.B. (1985). Understanding correlates of change by modeling individual differences in growth. Psychometrika, 50, 203-228.

Rosenbaum, P.R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. Journal of the Royal Statistical Society, Series A, 147, 656-666.

Rosenbaum, P.R. & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41-55.

Rosier, P. & Farella, M. (1976). Bilingual education at Rock Point, some early results. TESOL Quarterly, 10, 379-388.

Rumberger, R. (1983). Dropping out of high school: The influence of race, sex, and family background. American Educational Research Journal, 20, 199-220.

SAS Institute Inc. (1985). SAS User's Guide: Statistics, Version 5 Edition. Cary, NC: SAS Institute Inc.

Schinke-Llano, L., & Ramirez, J.D. (1984). Programmatic and Instructional Aspects of Language Immersion Programs. Mountain View, CA: SRA Technologies, Inc.

Skutnabb-Kangas, T. (1981). Bilingualism or Not: The Education of Minorities. Clevedon, Avon: Multilingual Matters.

Swain, M., Lapkin, S., Rowen, N., & Hart, D. (in press). The role of mother tongue literacy in third language learning. Language, Culture and Curriculum.

Twyford, C. (1988). Age-related factors in second language acquisition. New Focus, 2, 8-13.

United States General Accounting Office (1987). School Dropouts: Survey of Local Programs. Washington, DC: United States Government Printing Office.

Vargas, A. (1988). Literacy in the Hispanic Community. Washington, DC: National Council of La Raza.

Willett, J.B. (1989). Questions and answers in the measurement of change. Review of Research in Education. Washington: American Educational Research Association.

APPENDIX A

Number of Study Schools by District and Program

Table A-1

Number of Study Schools by District and Program

District	School	Program	District	School	Program
A	01	IS	F	50	EE
	02	EE		51	IS/EE
	04	EE			
	05	EE	G	60	LE
				61	LE
B	10	IS		62	LE
	11	EE		63	LE
	12	IS/EE		64	LE
	13	EE		65	LE
	14	IS/EE			
	15	EE	H	71	IS
				72	IS
C	20	EE		73	IS
	21	IS/EE		74	IS
	22	IS		75	IS
	23	EE		76	IS
	24	IS		77	IS
				78	IS
D	30	LE		79	IS
				91	IS
E	40	LE		92	IS
	41	LE		93	IS
	42	LE			
	43	LE	I	81	EE
	44	LE		82	EE
	45	LE		83	EE
	46	LE		84	EE

APPENDIX B

List of Covariates by Label and Name

Table B-1

List of Covariates by Label and Name

ABSENT	Days absent per year (specific to year of test score)
ABS1	Days absent in first grade
AGEMOTH	Age of the student's mother, in 3 categories
ANALPREL	Language pretest
ANALPREM	Mathematics pretest
ANALTOBE	Sum of mathematics and language pretest scores i.e., TOBE
ANYECTOP	Whether parents reported their children using English when talking to the parents
ANYEPTOC	Whether parents reported using English when talking to their children
ANYEPTOP	Whether parents reported using English when talking to each other
AVABS	Average number of absences per year
AVABS13	Average number of absences per year for grades one through three
BASE	Overall average
BOOKSHM	Number of books in the home, in 5 categories
BOOKSRD	Number of books parents have read in last 3 months
CAGEPGM	Child's age in months upon entering kindergarten
DISTA	Variable for district A
DISTB	Variable for district B
DISTC	Variable for district C
DISTF	Variable for district F
DISTG	Variable for district G
DISTHI	Variable for combination of districts H and I
EDAVG	Average years of education of student's parents

Table B-1
(Continued)

EGRADE3	Student is in late-exit 3-6 cohort
EE-A	Student is in early-exit program in district A
EE-B	Student is in early-exit program in district B
EE-F	Student is in early-exit program in district F
EE-HI	Student is in early-exit program in district I
E0SCHL44	Student is in K-3 cohort in school number 44
E3DISTG	Student is in 3-6 cohort in district G
FEMALE	Student was a girl
NEE02	School level variable for early-exit school number 02
NEE04	School level variable for early-exit school number 04
NEE05	School level variable for early-exit school number 05
NEE11	School level variable for early-exit school number 11
NEE13	School level variable for early-exit school number 13
NEE15	School level variable for early-exit school number 15
NEE50	School level variable for early-exit school number 15
NEE81	School level variable for early-exit school number 81
NEE82	School level variable for early-exit school number 82
NEE83	School level variable for early-exit school number 83
NEE84	School level variable for early-exit school number 84
NIS01	School level variable for immersion strategy school number 01
NIS10	School level variable for immersion strategy school number 10
NIS7A	School level variable for immersion strategy school group 7A
NIS7B	School level variable for immersion strategy school group 7B
NIS7C	School level variable for immersion strategy school group 7C

Table B-1
(Continued)

NIS7D	School level variable for immersion strategy school group 7D
OCAVG	Average SES occupation index for parents' occupations
PGMIS	Student is in immersion strategy program
PONEK1	Propensity score for students in K-1 analyses without sum of pretest score as a predictor
PONEK1PN	Propensity score for students in K-1 analyses who have a pretest score available without sum of pretest score as a predictor
PONEK1PP	Propensity score for students in K-1 analyses who have a pretest score available with sum of pretest score as a predictor
PONE13	Propensity score for students in 1-3 analyses without sum of pretest score as a predictor
PONE13PN	Propensity score for students in 1-3 analyses who have a pretest score available without sum of pretest score as a predictor
PONE13PP	Propensity score for students in 1-3 analyses who have a pretest score available with sum of pretest score as a predictor
PRESCHY	Whether student attended preschool
RPTOCLSY	Parent or someone else reads to child in Spanish
SCHOOL12	Student is in immersion strategy/early-exit school number 12
SCHOOL14	Student is in immersion strategy/early-exit school number 14
SCHOOL30	Student is in late-exit school number 30 (only school in district D)
SCHOOL45	Student is in late-exit school number 45
SCHOOL51	Student is in immersion strategy/early-exit school number 51

APPENDIX C

Average Achievement Scores

(Unadjusted and Adjusted)

ERIC

APPENDIX C

AVERAGE ACHIEVEMENT SCORES
(UNADJUSTED AND ADJUSTED)

Tables C-1 through C-12 give unadjusted means and standard deviations of the CTBS achievement test scores for each of the three subtests, separately for each of the three major divisions of the analyses. Tables C-1 through C-3 give the means and standard deviations by program and grade for the students in two-program schools (Chapter III presents the analyses of these students). Tables C-4 through C-6 give the means and standard deviations by program and grade for the students in one-program schools (Chapter IV presents the analyses of these students). Tables C-7 through C-12 give the means and standard deviations by district and grade for the students in the K-3 and 3-6 cohorts in the late-exit program (Chapter V presents the analyses of these students).

Tables C-13 through C-24 give the corresponding means after adjusting for the variables found to predict achievement. They are derived from the basic models presented in Chapters III through V by calculating the predicted value for hypothetical students. The hypothetical students are assumed to have covariate values exactly equal to the average over all 1164 students included in the analyses (see Chapter II for more information on the covariates, including the averages and standard deviations). However, the number of absences is assumed to be zero. Thus the tables of adjusted means show the predicted achievement for an "average" student with no absences.

Tables C-13 through C-15 correspond to the two-program schools 1-3 analyses in Chapter III. The SCHOOL variables are treated as zero for purposes of calculating the adjusted means, so that the adjusted mean reflects the average of the four two-program schools.

Tables C-16 through C-18 correspond to the basic two-level 1-3 analyses for one-program schools in Chapter IV. The NIS and NEE variables are treated as zero for purposes of calculating the adjusted means, so that the adjusted mean for each program reflects the average of the school groups with that program. Note that the program difference at e ch grade incorporates the curvature difference.

Tables C-19 through C-24 correspond to the late-exit 1-6 analyses excluding the K-3 students in district G, from Chapter V. The adjusted means for the K-3 and 3-6 cohorts are presented in separate tables. The five columns of adjusted means reflect the school and district differences found in the analyses: for some subtests, school 30 (the only school in district D), school 44, school 45, or the schools in district G as a group differ significantly from schools 41, 42, 43, and 46 in district E. For ease of comparison, the five separate columns are presented even for the models that include no difference between pairs of columns. For example, Table C-19 for the K-3 cohort shows the predicted mathematics achievement score to be the same in school 30 in district D as for the group of four schools in district E (schools 41, 42, 43, and 46) for all three grades.

Table C-1

Unadjusted
Average Mathematics CTBS Achievement Scores by Program and Grade
Two-Program Schools

	Immersion Strategy		Early-Exit	
Grade	Mean	S.D.	Mean	S.D.
1	245.468	25.552	251.925	31.982
2	290.043	23.861	293.939	30.271
3	323.674	34.307	332.744	34.066

Table C-2

Unadjusted
Average English Language CTBS Achievement Scores by Program and Grade
Two-Program Schools

	Immersion Strategy		Early-Exit	
Grade	Mean	S.D.	Mean	S.D.
1	239.424	35.718	231.881	34.291
2	310.804	47.096	298.242	48.383
3	361.054	55.310	364.581	56.622

Table C-3

Unadjusted
Average English Reading CTBS Achievement Scores by Program and Grade
Two-Program Schools

	Immersion Strategy		Early-Exit	
Grade	Mean	S.D.	Mean	S.D.
1	220.820	30.437	225.985	25.839
2	296.355	36.510	286.394	42.481
3	333.587	38.376	339.744	42.912

Table C-4

Unadjusted
Average Mathematics CTBS Achievement Scores by Program and Grade
One-Program Schools

	Immersion Strategy		Early-Exit	
Grade	Mean	S.D.	Mean	S.D.
1	253.649	31.535	253.579	32.412
2	304.870	33.947	301.580	34.237
3	344.259	41.335	351.154	42.644

Table C-5

Unadjusted
Average English Language CTBS Achievement Scores by Program and Grade
One-Program Schools

	Immersion Strategy		Early-Exit	
Grade	Mean	S.D.	Mean	S.D.
1	261.196	46.328	256.722	50.148
2	347.413	62.493	321.428	64.762
3	390.071	74.856	384.994	69.634

Table C-6

Unadjusted
Average English Reading CTBS Achievement Scores by Program and Grade
One-Program Schools

	Immersion Strategy		Early-Exit	
Grade	Mean	S.D.	Mean	S.D.
1	236.289	39.806	240.167	39.854
2	306.739	54.925	301.492	51.023
3	351.143	57.936	352.101	54.115

Table C-7

Unadjusted
Average Mathematics CTBS Achievement Scores by District and Grade
Late-Exit K-3 Cohort

	Late-Exit Site D		Late-Exit Site E		Late-Exit Site G	
Grade	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	269.600	27.457	242.035	24.098	*	*
2	301.097	39.678	294.859	34.031	303.469	36.806
3	384.871	38.584	336.900	53.439	336.735	43.819

* = No data available at this grade level for this site.

Table C-8

Unadjusted
Average English Language CTBS Achievement Scores by District and Grade
Late-Exit K-3 Cohort

	Late-Exit Site D		Late-Exit Site E		Late-Exit Site G	
Grade	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	310.743	49.021	241.430	31.693	*	*
2	356.839	53.741	280.847	58.196	293.265	71.888
3	443.258	66.398	331.025	67.106	353.796	55.688

* = No data available at this grade level for this site.

Table C-9

Unadjusted
Average English Reading CTBS Achievement Scores by District and Grade
Late-Exit K-3 Cohort

	Late-Exit Site D		Late-Exit Site E		Late-Exit Site G	
Grade	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	274.686	46.235	224.919	25.590	*	*
2	333.806	52.698	274.647	46.027	294.224	47.013
3	401.290	62.868	327.975	57.429	342.612	49.893

* = No data available at this grade level for this site.

Table C-10

Unadjusted
Average Mathematics CTBS Achievement Scores by District and Grade
Late-Exit 3-6 Cohort

	Late-Exit Site D		Late-Exit Site E		Late-Exit Site G	
Grade	Mean	S.D.	Mean	S.D.	Mean	S.D.
3	402.000	60.726	347.832	43.734	354.923	41.314
4	435.238	48.354	372.573	44.706	380.214	40.307
5	406.333	88.741	419.453	55.055	418.100	34.859
6	492.059	98.916	460.591	67.497	388.381	52.037

Table C-11

Unadjusted
Average English Language CTBS Achievement Scores by District and Grade
Late-Exit 3-6 Cohort

	Late-Exit Site D		Late-Exit Site E		Late-Exit Site G	
Grade	Mean	S.D.	Mean	S.D.	Mean	S.D.
3	427.450	63.586	345.589	46.085	357.500	65.779
4	456.381	64.061	378.094	44.022	392.357	55.561
5	463.429	88.392	404.427	46.535	413.950	31.673
6	508.824	92.254	441.364	48.702	422.571	46.907

Table C-12

Unadjusted
Average English Reading CTBS Achievement Scores by District and Grade
Late-Exit 3-6 Cohort

	Late-Exit Site D		Late-Exit Site E		Late-Exit Site G	
Grade	Mean	S.D.	Mean	S.D.	Mean	S.D.
3	403.050	71.255	342.316	51.936	319.615	64.182
4	433.238	61.282	355.510	54.287	378.929	60.076
5	440.524	63.938	415.800	54.306	403.500	32.180
6	500.000	69.531	454.000	46.870	406.667	57.725

Table C-13

Adjusted
Average Mathematics CTBS Achievement Scores by Program and Grade
Two-Program Schools

Grade	Immersion Strategy	Early-Exit
1	249	255
2	297	301
3	331	334

Table C-14

Adjusted
Average English Language CTBS Achievement Scores by Program and Grade
Two-Program Schools

Grade	Immersion Strategy	Early-Exit
1	249	238
2	313	307
3	357	355

Table C-15

Adjusted
Average English Reading CTBS Achievement Scores by Program and Grade
Two-Program Schools

Grade	Immersion Strategy	Early-Exit
1	228	232
2	301	290
3	330	338

Table C-16

Adjusted
Average Mathematics CTBS Achievement Scores by Program and Grade
One-Program Schools

Grade	Immersion Strategy	Early-Exit
1	258	256
2	308	305
3	348	352

ERIC

Table C-17

Adjusted
Average English Language CTBS Achievement Scores by Program and Grade
One-Program Schools

Grade	Immersion Strategy	Early-Exit
1	256	257
2	342	324
3	393	384

Table C-18

Adjusted
Average English Reading CTBS Achievement Scores by Program and Grade
One-Program Schools

Grade	Immersion Strategy	Early-Exit
1	233	240
2	304	304
3	353	354

ERIC

Table C-19

Adjusted
Average Mathematics CTBS Achievement Scores by District/School and Grade
Late-Exit K-3 Cohort

Grade	District D	District E (Schl 44)	District E (Schl 45)	District E (Schls 41-43,46)	District G
1	253	262	237	253	*
2	305	295	289	305	*
3	357	328	341	357	*

* = No data available for these grade levels for this site.

Table C-20

Adjusted
Average English Language CTBS Achievement Scores
by District/School and Grade
Late-Exit K-3 Cohort

Grade	District D	District E (Schl 44)	District E (Schl 45)	District E (Schls 41-43,46)	District G
1	290	262	239	239	*
2	347	275	296	296	*
3	403	288	353	353	*

* = No data available for these grade levels for this site.

Table C-21

Adjusted
Average English Reading CTBS Achievement Scores
by District/School and Grade
Late-Exit K-3 Cohort

Grade	District D	District E (Schl 44)	District E (Schl 45)	District E (Schls 41-43,46)	District G
1	262	242	227	227	*
2	321	269	286	286	*
3	379	295	344	344	*

* = No data available for these grade levels for this site.

Table C-22

Adjusted
Average Mathematics CTBS Achievement Scores by District/School and Grade
Late-Exit 3-6 Cohort

Grade	District D	District E (Schl 44)	District E (Schl 45)	District E (Schls 41-43,46)	District G
3	366	366	350	366	365
4	394	394	378	394	378
5	422	422	406	422	392
6	450	450	434	450	405

Table C-23

Adjusted
Average English Language CTBS Achievement Scores
by District/School and Grade
Late-Exit 3-6 Cohort

Grade	District D	District E (Schl 44)	District E (Schl 45)	District E (Schls 41-43,46)	District G
3	410	359	359	359	362
4	434	384	384	384	385
5	459	408	408	408	409
6	483	433	433	433	433

Table C-24

Adjusted
Average English Reading CTBS Achievement Scores
by District/School and Grade
Late-Exit 3-6 Cohort

Grade	District D	District E (Schl 44)	District E (Schl 45)	District E (Schls 41-43,46)	District G
3	383	348	348	348	335
4	414	379	379	379	364
5	444	409	409	409	392
6	475	440	440	440	421

ERIC