ED 363 733   CE 064 961

AUTHOR	Venezky, Richard L.; And Others
TITLE	When Less Is More: A Comparative Analysis for Placing Students in Adult Literacy Classes. TR93-8.
INSTITUTION	National Center on Adult Literacy, Philadelphia, PA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	NCAL-TR-93-8
PUB DATE	Sep 93
CONTRACT	R117Q0003
NOTE	41p.
AVAILABLE FROM	National Center on Adult Literacy, Dissemination/Publications, University of Pennsylvania, 3910 Chestnut Street, Philadelphia, PA 19104-3111 ($5; checks payable to Kinko's Copy Center).
PUB TYPE	Reports - Research/Technical (143)

EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Adult Basic Education; Adult Literacy; Adult Programs; Educational Testing; *High School Equivalency Programs; *Literacy Education; *Norm Referenced Tests; Predictive Validity; *Student Placement; Test Validity
IDENTIFIERS	General Educational Development Tests; *Tests of Adult Basic Education; *Tests of Adult Literacy Skills

ABSTRACT

A study evaluated different approaches to placing students in adult literacy classes using 168 adult basic education (ABE) and General Educational Development (GED) students. Students in classes at Rochambeau School in White Plains, New York, were given the Tests of Adult Basic Education (TABE), the Tests of Adult Literacy Skills (TALS), an oral reading test, and a decoding test. They were tested upon entry into an adult literacy program and then placed into either ABE classes (three levels) or a GED preparation class (one level). Various mismatch measures, along with multiple and linear regression, were used to evaluate how well each test predicted the actual placements made. The TABE Locator, which required 37 minutes to administer, predicted actual placements as well as or better than the full TABE, which was normally administered for placement and which, with the Locator, required almost 3 hours to administer. The conclusion from these data was that lengthy testing for placement into four or fewer levels is unnecessary. Neither the TALS, the decoding test, nor the oral reading test performed as well as the TABE Locator. (Appendixes include 24 references, 5 data figures, descriptions of White Plains placement procedures, and 6 data tables.) (Author/YLB)

NCAL

NATIONAL CENTER ON ADULT LITERACY

WHEN LESS IS MORE:

A COMPARATIVE ANALYSIS FOR PLACING STUDENTS IN ADULT LITERACY CLASSES

Richard L. Venezky
Page S. Bristow
John P. Sabatini
University of Delaware

**NCAL TECHNICAL REPORT TR93-8
SEPTEMBER 1993**

NATIONAL CENTER ON ADULT LITERACY
UNIVERSITY OF PENNSYLVANIA

2

WHEN LESS IS MORE:

A COMPARATIVE ANALYSIS FOR PLACING STUDENTS IN ADULT LITERACY CLASSES

Richard L. Venezky
Page S. Bristow
John P. Sabatini
University of Delaware

**NCAL TECHNICAL REPORT TR93-8
SEPTEMBER 1993**

The National Center on Adult Literacy (NCAL) was established in 1990 by the U.S. Department of Education, with co-funding from the Departments of Labor and Health and Human Services. The mission of NCAL addresses three primary challenges: (1) to enhance the knowledge base about adult literacy, (2) to improve the quality of research and development in the field and (3) to ensure a strong, two-way relationship between research and practice. Through applied research and development and dissemination of the results to researchers, policymakers and practitioners, NCAL seeks to improve the quality of adult literacy programs and services on a nationwide basis. NCAL serves as a major operating unit of the Literacy Research Center at the University of Pennsylvania.

NCAL publications to date include:

May 1992   *Matching Literacy Testing With Social Policy: What are the Alternatives?*
Richard L. Venezky   (PB92-1, 8 pages)

Oct 1992   *Life-Span and Life-Space Literacy: Research and Policy in National and International Perspectives*
Daniel A. Wagner   (OP92-1, 24 pages)

Oct 1992   *Expanding Theories of Adult Literacy Participation*
Karen Reed Wikelund, Stephen Reder, Sylvia Hart-Landsberg   (TR92-1, 40 pages)

Oct 1992   *Invitations to Inquiry: Rethinking Staff Development in Adult Literacy Education*
Susan L. Lytle, Alisa Belzer, Rebecca Reumann   (TR92-2, 44 pages)

Dec 1992   *Developing the Professional Workforce for Adult Literacy Education*
Susan L. Lytle, Alisa Belzer, Rebecca Reumann   (PB92-2, 11 pages)

Jan 1993   *The Impact of BIB-Spiralling Induced Missing Data Patterns on Goodness-of-Fit Tests in Factor Analysis*
David Kaplan   (OP93-1, 18 pages)

Mar 1993   *The Impact of Workplace Literacy Programs: A New Model for Evaluation of Workplace Literacy Programs*
Larry Mikulecky, Paul Lloyd   (TR93-2, 180 pages)

Mar 1993   *Literacy and Machines: An Overview of the Use of Technology in Adult Literacy Programs*
Terilyn C. Turner   (TR93-3, 86 pages)

Jun 1993   *Literacy and Development: Rationales, Assessment, and Innovation*
Daniel A. Wagner   (IP93-1, 50 pages)

Jun 1993   *Myths and Misconceptions in Adult Literacy: A Research and Development Perspective*
Daniel A. Wagner   (PB93-1, 10 pages)

Jun 1993   *Early Childhood, Family, and Health Issues in Literacy: International Perspectives*
Laurei D. Puchner   (IP93-2, 45 pages)

Sep 1993   *Prison Literacy: Implications for Program and Assessment Policy*
Anabel Newman, Warren Lewis, Carolyn Beverstock   (TR93-1, 219 pages)

Sep 1993   *Management Information Systems in Adult Education: Perspectives from the States and from Local Programs*
Mark A. Kutner, Lenore Webb, Rebecca Herman, Pelavin Associates, Inc.   (TR93-4, 150 pages)

4

NCAL publications to date (continued)

Information on ordering of NCAL publications may be addressed to Dissemination at NCAL.

September 10, 1993

TABLE OF CONTENTS

A Joint Project of the New York State Education Department, the University of Delaware, and the Adult and Continuing Education Program of the White Plains, NY Public Schools.

7

WHEN LESS IS MORE:

A COMPARATIVE ANALYSIS FOR PLACING STUDENTS IN ADULT LITERACY CLASSES

Richard L. Venezky
Page S. Bristow
John P. Sabatini
University of Delaware

Abstract

This study evaluated different approaches to placing students in adult literacy classes using 168 Adult Basic Education (ABE) and General Equivalency Degree (GED) students who were given the Tests of Adult Basic Skills (TABE), the Tests of Adult Literacy Skills (TALS), an oral reading test, and a decoding test. Students were tested upon entry into an adult literacy program and then placed into either ABE classes (three levels) or a GED preparation class (one level). Various mismatch measures, along with multiple and linear regression, were used to evaluate how well each test predicted the actual placements made. The TABE Locator, which requires 37 minutes to administer, predicted actual placements as well as or better than the full TABE, which are normally administered for placement and which, with the Locator, require almost three hours to administer. From these data, it was concluded that lengthy testing for placement into four or fewer levels is unnecessary. By reducing placement testing, literacy programs will not only have more time for instruction, but may also increase retention of many students who might be put off by more complex standardized tests.

8

INTRODUCTION

One of the most important uses of norm-referenced tests in adult literacy programs is for initial placement into instructional levels. In a study by Ehringhaus (1991), almost 85% of those responding to the question "Why do programs use formal tests?" marked "usually" or "always" for "To begin students at the right level." The federal government currently funds two forms of adult literacy programs: Adult Basic Education (ABE), commonly divided into three levels referred to as ABE 1, ABE 2, and ABE 3, and General Equivalency Degree (GED) preparation. At larger adult literacy sites, separate classes are held during the day and in the evening for all four levels. However, most sites do not have the attendance or the staffing required for this full complement of courses; nonetheless, students are usually placed into one of these levels by one mechanism or another, even though a teacher may have students who range across two or more adult literacy levels in a single classroom.

Placement receives little attention in the testing literature, aside from admissions and placement issues at the college and university level (Whitney, 1989; Willingham, 1974). In most practical placement decisions, individuals must be divided among a fixed number of levels in a hierarchical sequence (beginning, intermediate, or advanced French, for example) or into discrete categories, such as military occupational specialties. In the latter case, multiple discriminant functions can be used to match score profiles of new persons to be placed to those of persons already in each group or specialty (Nunnally, 1967). In the former case, which is relevant to the placement of students into adult literacy classes, a single test is commonly used, with cut scores established for placement into each available instructional level. When student abilities are known from prior instructional experience, testing is not necessary (Nitko, 1989).

Ideally, a person should be placed at an instructional level where he or she has met all of the important prerequisites but none of the important instructional outcomes (Whitney, 1989). This ideal cannot be met for ABE and GED classes because no practical program can offer as many classes as would be needed to meet the wide variety of student entry-level abilities typically encountered. An ABE program could require as many as 81 separate classes, for example, to teach reading, writing, mathematics, and life skills, assuming that students were divided

9

into three entry levels on each of these skill areas. Therefore, programs typically offer no more than three levels of ABE classes plus a GED class, and placement is often based on reading ability alone or on a combination of reading and mathematics abilities.

The most common approach to placement is the use of a battery of tests such as the Tests of Adult Basic Education (TABE) (CTB/McGraw-Hill, 1987b). In a recent nationwide survey of 427 ABE teachers, the TABE was mentioned by 59% as the only test regularly used or as one of several tests in regular usage. The nearest competitor was the Adult Basic and Literacy Education (ABLE), mentioned by fewer that 5% of the respondents (Ehringhaus, 1991). The National Evaluation of Adult Education Programs reported that the TABE is used for initial diagnosis of learner needs by 68% of the 2,619 programs responding to their survey (Development Associates, 1992).

Typically, the TABE Locator Test, consisting of separate, 25-item multiple choice tests for vocabulary and mathematics computation, is given to determine which of the four overlapping levels (easy, medium, difficult, advanced) of the TABE Tests should be administered. Some programs administer both the Reading Comprehension and Vocabulary Tests, along with the Mathematics Computation and Mathematics Concepts and Applications Tests. All together, almost three hours are required to administer these tests according to the publisher's guidelines (CTB/McGraw-Hill, 1987a). Additional time is required for scoring and for determining placement, which is generally done using tables supplied by the test publisher.

A number of problems have surfaced repeatedly in the placement procedure just outlined. First of all, many individuals who work up the courage to attend an adult literacy program are discouraged or intimidated by this long testing procedure, especially since it typically occurs within the first few days of enrollment. Many of these people have had bad experiences in formal schooling and have not seen an academic skills test for many years. Testing at program entry is a strong disincentive to continue in a program (Lytle & Wolfe, 1989).

Second, the reliability of tests given to individuals returning to formal education after five or more years may not be as high on days one and two of instruction as it might be after a few weeks or months of instruction (Sticht, 1990). Although there have been no studies that address this issue directly, the larger study from which the data reported here derive reveals different test-taking patterns for adults tested at different times during the academic year.

During the first instance of testing, given at the time of initial enrollment, students tended to work tests from beginning to end, attempting to answer every question. This may have led to perseverance on one or more items, resulting in omissions mainly at the end of the test. During subsequent testings, all coming after dozens of hours of instruction, better test-taking strategies were evident from the scatter of omissions across the tests, indicating that the students were skipping over items they thought were too difficult. This strategy also resulted in more items attempted, an advantage on a multiple-choice test that has no penalty for wrong answers.

Finally, with only four levels of classes in which to place students, the necessity of hours of formal testing, with its concomitant costs to both students and staff, should be questioned. For the TABE in particular, the Locator Test alone might function as an adequate placement instrument, as might a brief oral reading assessment or a decoding assessment.

To investigate this issue, a study was conducted on the placement predictions that could be drawn from a variety of different assessment procedures, including the TABE Locator, the TABE Tests, the relatively new Tests of Adult Literacy Skills (TALS), an oral reading test, and a decoding test. This study was guided by two rough hypotheses. First, it was suspected that a brief test like the TABE Locator would function as well as any of the full-length TABE tests for the purpose of placement; and second, it was predicted that the oral reading and decoding tests would be as accurate an indicator of placement as any of the full-length TABE or TALS Tests.

11

A. METHODS

1. SUBJECTS

All subjects for this study attended ABE or GED classes at the Rochambeau School in White Plains, New York, during the 1991-92 school year. This school is the site of the White Plains Adult and Continuing Education Program and is used exclusively for that purpose. Besides the ABE and GED classes, the school offers an extensive number of programs, including English-as-a-Second-Language (ESL), job skills, general continuing education, workplace literacy, neighborhood literacy, and family literacy, in cooperation with White Plains elementary schools. Many of the subjects for this project were graduates of the school's ESL programs. All of the students attended either ABE 1, ABE 2, ABE 3, or GED classes, either during the day or in the evening. Of the 213 students who registered for classes, 168 completed all of the full-length tests of the TABE and the TALS.

The subjects were predominantly foreign-born, non-Caucasian, low-income, and either unmarried or separated from their spouses. There were slightly more males (53%) than females (47%), and almost 60% were in the age range 26-50. Few of the subjects had voted during the past five years in a national or state election, almost none reported any health-related handicaps, one-quarter read a newspaper daily, and nearly three-quarters considered themselves sufficiently literate to handle the reading demands of home, work and family. Most also claimed to have relatively extensive literacy practices, as evidenced by self-reports of newspaper, magazine, book, and other types of reading. All were attending adult literacy classes voluntarily and most were pursuing improved literacy as a step towards occupational advancement. Figures 1 to 4 (see Appendix A) profile some of the background characteristics of this population.

2. DESIGN

This placement study was part of a larger study, conducted in cooperation with the New York State Department of Education and the White Plains Public Schools, aimed at the improvement of adult literacy assessment. Selection of the TABE and the TALS resulted from the interest of the cooperating organizations in comparing these tests for a variety of adult literacy assessment needs. For the placement component, the White Plains program followed its normal placement procedure (see Appendix A), but

with a collapsing of levels of the TABE that were administered. In addition, other tests were administered at the same time, as described below. The purpose of this aspect of the study was to determine the degree to which each assessment procedure could predict the placements that resulted from the normal placement procedure, which for lack of any external validation was accepted as optimal.[1] A listing of the tests administered is given in Table 1 (see Appendix C).

3. INSTRUMENTS

A. TESTS OF APPLIED LITERACY SKILLS (TALS)

The TALS is a battery of norm-referenced tests that use functional literacy tasks to measure an adult's ability to apply literacy skills in contexts commonly encountered in everyday living. These instruments were developed from the experiences gained by the Educational Testing Service with the Young Adult and Department of Labor Literacy Surveys (Kirsch & Jungeblut, 1986; Kirsch, Jungeblut, & Campbell, 1992). TALS items require short answer and other constructed responses as opposed to the multiple-choice format of other tests.

The TALS battery is composed of three tests: Document Literacy, Prose Literacy, and Quantitative Literacy. However, only the Document and Quantitative Literacy Tests were included in this study. According to the TALS publisher, the Document Literacy Test (DL) measures the ability to identify and use information located in materials such as charts, indices, forms, and tables. The Quantitative Literacy Test (QL) requires the performance of arithmetic operations, alone or sequentially, using various printed materials; for example, respondents calculate wages, complete an order form, and determine the amount of interest from a loan advertisement. Each test is divided into two sections that are scored separately (Kirsch, Jungeblut, & Campbell, 1991) using the scoring guide provided by the publisher. The two section scores then serve as indices to retrieve a single scale score from a publisher-supplied table. Scale scores were derived through IRT scaling from a norming study that involved 3,105 adults (Simon & Schuster, 1990). Each test has two alternate forms and is administered in a 40-minute period.

B. TESTS OF ADULT BASIC EDUCATION (TABE)

The TABE is a battery of norm-referenced tests that require multiple-choice responses. The tests administered in this study were the Vocabulary, Reading Comprehension, Mathematics Computation, and Mathematics Concepts and Applications Tests.

_13_

According to the publisher, the purpose of the battery is not to test specific life skills, but to test basic skills in the context of life skill tasks (CTB/McGraw-Hill, 1987c). The TABE Vocabulary Test measures mastery of synonyms, antonyms, homonyms, affixes, and words in context. The Reading Comprehension Test measures literal, inferential and critical comprehension. The Mathematics Computation Test measures the ability to do addition, subtraction, multiplication, and division. The Mathematics Concepts and Applications Test measures numerical concepts such as place-value, number sentences, and geometry, as well as reasoning skills needed for practical problem-solving. Each test has four graduated but overlapping levels, (easy, medium, difficult, advanced) with alternate forms available for each. Also available is a Locator Test for determining the appropriate level for full-scale testing. This Locator includes 25 multiple-choice vocabulary items and 25 multiple-choice arithmetic items and requires 37 minutes for administration.

Test scores are converted from table values to scale scores derived from IRT scaling. Norming of the tests was done with about 6,300 examinees, divided among adult basic education enrollees, adult offenders, juvenile offenders, and vocational/technical school enrollees. As Figure 5 shows (see Appendix A), the scale ranges for the four levels have considerable overlap.

C. ORAL READING TASKS PASSAGES

Although adults engage in silent reading far more often than they do oral reading, studies of low-literate adults have shown that oral reading ability is a consistent indicator of reading comprehension (Bristow & Leslie, 1988). Oral reading rate might, therefore, be an adequate indicator of reading level. To test this possibility, four passages for oral reading were selected from a variety of instructional materials commonly used in ABE and GED programs. Each passage was selected from expository materials of ascending difficulty and was minimally adapted for length. The resulting passages varied from 188 to 328 words in length. The topics (sleepwalkers, lightning, plastic trash, and fever) were selected because of their high familiarity among virtually all adult populations. Two comprehension questions were prepared for each passage. One question required factual recall and the other required an inference based upon textual material. The questions were administered solely to focus the adults' attention on comprehension as they read orally.

14

D. ORAL READING TASKS: DECODING

The decoding tasks consisted of seven lists of six pseudo-words each, designed to be of increasing difficulty. List one consisted of three-letter consonant-vowel-consonant pseudo-words. List two consisted of four-letter pseudo-words, with consonant clusters and short vowels. List three included four- and five-letter pseudo-words with variant consonant pronunciations, digraphs, long vowels and silent-letter Es. List four contained pseudo-words with vowel combinations (digraphs). List five contained some multisyllabic pseudo-words and more complex vowel and consonant combinations. List six contained two- and three-syllable pseudo-words, whereas list seven contained four- and five-syllable pseudo-words composed of high frequency syllables. A summary of these lists, along with sample test items, is shown in Table 2 (see Appendix C). Internal consistency of the entire test, as measured by Kuder-Richardson Formula 20 (KR-20), varied from 69% to 81.2% over the various testing periods incorporated into the larger study.

E. BACKGROUND QUESTIONNAIRE

The Background Questionnaire, adapted from the questionnaire developed and used by ETS in the National Adult Literacy Survey (NALS), is composed of six sections: general and language background, educational background and experiences, political and social participation, labor force participation, literacy activities, and demographic information. The questionnaire was administered individually in an interview format and required 15 to 20 minutes for completion.

4. PROCEDURES

A. TABE/TALS ADMINISTRATION

Students were randomly assigned to take either the TABE or the TALS on day one; the remaining tests were given on the next class day. Each set of tests was administered in a single sitting; group administration in classrooms utilized the publisher's standardized instructions, including time limits. Students were placed into one of two levels of the TABE (easy or difficult), based on their total (vocabulary plus mathematics) TABE Locator Test score. This reduction in levels was done both to reduce the complexity of the study design and to increase the power of several of the tests projected for the larger study. Students who scored less than 12 on the Locator Test were considered non-readers and thus did not take the TABE (or TALS) battery. Students who received raw scores between 13 and 29 were given the E (easy) level, and subjects who

scored above 39 were given the D (difficult) level. Subjects whose scores were between 30 and 39 were randomly assigned to either the D or the E levels. Normally, these students would have been placed in Level M (medium), but since the tests overlap considerably in difficulty levels, little loss in precision was projected. All test administrators attended a three-hour training session that prepared them to use the TABE and TALS standardized administration procedures and to administer the oral reading tasks as described below.

B. ORAL READING TASKS

Subjects were individually tested in a half-hour session that was audio taped. Twelve examiners participated in the testing; of the 12, 6 were ABE/GED teachers, 2 were guidance counselors, and 4 were college students who had prior experience working with the adult population being tested. The oral reading tasks included two sections: decoding and oral reading of passages. For the decoding section, subjects were told that they would be shown lists of made-up words that they were to read out loud. Prior to reading the lists, students read a sample to give them practice with made-up (pseudo) words. Once they appeared to understand the task, the lists were administered. The made-up words were presented on cards that displayed six words each. Each student read the seven lists in order with no interruptions unless he or she began to have difficulty. If the student was obviously struggling, the examiner asked, "Would you like to stop?" The student then made the decision to stop or continue.

Regardless of whether the student completed the decoding lists. he or she was asked to read the oral reading passages. Students read orally as many of the four selections as they could and answered comprehension questions after each. The passages were ordered in ascending level of difficulty. Students were told that the examiner could not help them if they had difficulty and that after reading, they would be expected to answer questions without looking back at the passage. Comprehension questions were included to assure that the students were focusing on comprehension as a goal in reading. The instructions given here encourage use of what Carver (1990) calls the learning process. Carver found that the oral reading rate for individuals varies with the reading task. Skimming and scanning, for example, can be done significantly faster than *rauding,* Carver's term for reading comprehension of relatively easy materials. Learning and memorizing are the slowest forms of reading, according to Carver.

Subjects were told prior to reading that they could stop at any point if the reading became too difficult for them. Forty-seven students exercised this option before completing all four passages. Data on passages begun but not completed were excluded from the study. After reading each passage, subjects answered two comprehension questions and made a self-assessment of the difficulty of the passage. Subjects were asked to rate the passage as easy, hard, or just about right for them. Because these data do not apply directly to the placement analysis, they are not reported here.

C. BACKGROUND QUESTIONNAIRE

Each background questionnaire was administered in interview format by one of eight examiners who had attended a 3-hour training session on administration procedures. For all questions with multiple-choice responses, hand cards were presented which listed the potential responses. The choices were read to the student from the card, after which he or she selected one of the alternatives (or supplied a new one). Students were notified at the beginning of the interview that they could refuse to answer any question.

5. SCORING

A. ORT DECODING/PASSAGES

The ORT decoding lists were scored by a linguistics doctoral student trained by the author who designed the decoding task. A random sample of decoding protocols was rescored by the author to verify that scoring procedures were being followed systematically. A number of issues were uncovered in the scoring procedures, the most serious of which relates to the evaluation of non-native English pronunciation. In many cases, the distinction between incorrect decoding and a non-native pronunciation was difficult to make. Because of the high percentage of non-native English speakers in the main sample for this study, the reliability of the decoding test is somewhat suspect, but not so much as to lead to its removal from the study.

The ORT reading passages were timed from the audio tapes, using a stopwatch to determine minutes and seconds for the reading of each passage. All passages were retimed to verify the accuracy of the timing procedure. Agreement between the two timings was over 90%. Disagreements of more than 3 seconds were resolved by a third timing. Since correlations of reading rates across the passages for the entire study ranged from 0.85 to 0.98,

rates on a single passage (Lightning) were used for the analyses described below.

B. TABE/TALS

The TABE tests were scored twice, initially by test examiners and later by project personnel at the University of Delaware. Discrepancies were resolved by a third scoring. For the Locator Test, scoring errors made by the initial scorers totaled 11.8% for the vocabulary section and 11.3% for the mathematics section. Seventy-three percent of these errors were within two items of the correct score. The TALS tests were scored by an ETS-trained scorer, utilizing the standardized scoring criteria. Twenty percent of the TALS tests were rescored by another ETS-trained examiner; the interrater reliability was 99%.

18

B. RESULTS

In order to investigate placement predictions, four different types of analyses were conducted on the test results. Two involved mismatch computations for selected tests against actual placements; the third involved a multiple regression of various predictors of actual placements, and the fourth, separate linear regressions of selected tests. All four methods are strongly bound by several important limitations of the actual testing situation. First, there was no external criterion to validate that the actual placements made were in fact optimal for the choices of classes available (see Appendix A for a description of the placement procedures used in White Plains). Some shifting of students did occur after the initial placements, but this was minimal. In addition, in some classrooms, individualization of instruction was relatively high so that misplaced students might have nevertheless received appropriate instruction. Finally, the TABE was the only measure used by the program staff for placement; therefore, no unbiased assessment can be made of the TALS and oral reading tasks as placement measures.

In the first mismatch analysis, the complete scoring scales of the TABE Tests, the TALS Document Literacy and Quantitative Literacy Tests and the oral reading and decoding tests were divided into four equal ranges.[2] Under the assumption that an *ideal* placement algorithm would place everyone in the lower quarter of a scale into ABE 1, everyone in the next quarter into ABE 2, and so on, through the GED class, the absolute differences between predicted placements and actual placements for each test were summed. Thus, if a student were predicted by a particular test to be in ABE 2 but had actually been placed in a GED class, the mismatch would be 2. Similarly, a predicted placement of ABE 2, but actual placement in ABE 1, would yield a mismatch of 1.

Table 3 (see Appendix B) shows the mean mismatch and its standard deviation for each test. What is striking is that the TABE Locator Test, which requires 37 minutes to administer, was a better predictor of placement than any of the full TABE Tests, including Total Reading (a combination of the comprehension and vocabulary subtests). The Locator math and vocabulary subtests separately are nearly as effective for placement as the most effective full tests.

A second analysis of placement utilized Cohen's Kappa coefficient for nominal scales (Cohen, 1960). This test was

originally developed for determining agreement between pairs of judges or raters assigning subjects to nominal categories. Its primary virtues are (1) its control for agreement expected by chance and (2) its simplicity. Nevertheless, it is not a strong measure for educational class placements because classes are generally ordered along a difficulty scale and therefore do not represent true nominal scales. In the present case, Cohen's Kappa coefficient captures only the degree of complete agreement between a test and actual placement; what it fails to capture is the magnitude of mismatch for each subject. For example, a subject who is placed in a GED class but should be in ABE 1 represents a far more serious mismatch than an ABE 1 to ABE 2 mismatch, yet the Kappa coefficient treats the two equally. The Kappa coefficients for the various tests (see Table 4, Appendix C) show roughly the same rank ordering as the mean placement mismatches, although the intervals between tests differ. (The meaningful range for Kappa is 0-1.) Once again, the TABE Locator Test showed the highest match with actual overall placements but in this analysis was equaled by Total Reading.

The third analysis performed on placement data involved a multiple regression, using placement as a dependent measure and the various TABE and TALS tests and oral reading rate as predictors. (For this analysis, as with all the others, we assumed that the ABE/GED classes represented an equal-interval scale.) This test was run on 147 of the 168 subjects who completed all of the tests. An initial model constructed from the complete set of tests yielded a multiple correlation of 0.812 (adjusted $r^2 = 0.659$) with actual placements. From this model, predictors that had significant or nearly significant relationships to the dependent measure were retained and a second model created. This model yielded an adjusted r^2 with actual placement of 0.657 ($F=63.8$, $p<.001$), with the two Locator subtests having the highest ß-weights (See Tables 5 & 6, Appendix C). The TABE Comprehension and Mathematics Concepts and Applications Tests also made small but significant contributions to the regression. This outcome reinforces the mismatch analysis result.

Finally, separate linear regression models were run for the TABE Locator, TABE Reading Comprehension, TABE Total Reading (i.e., comprehension and vocabulary) and TALS Document Literacy, with actual placements as the dependent variable. These regressions provide independent measures of the predictive validities of each selected test. In agreement with the other analyses, the Locator had an adjusted squared multiple R value of 0.601 compared to 0.580 for TABE Total Reading, 0.511 for

Reading Comprehension and 0.281 for Document Literacy. (In all three cases the regression coefficients were highly significant: p < 0.000 for two-tailed t tests.) (See Tables 5 and 6, Appendix C.)

21

C. CONCLUSIONS

Within the bounds of the qualifications given in the Methods and Results sections, it can be concluded first that lengthy testing procedures are not necessary for placement. For programs that have no more than four distinguishable levels of instruction, a test as short as the TABE Locator Test is more effective for placement than any of the full TABE tests and at least as effective as the TABE Total Reading Comprehension and Vocabulary Tests combined. Neither the TALS, the decoding test, nor the oral reading test performed as well as the TABE Locator, although the design of the study did not allow a completely unbiased evaluation of these tests.

Thus, the first hypothesis, that a test as brief as the TABE Locator could predict placements at least as well as the complete group of reading tests, was confirmed. For many programs, the TABE Vocabulary or Mathematics Locator may be adequate, requiring only 15 to 19 minutes of testing time. The second hypothesis, however, was not supported. Neither the oral reading rate measure nor the decoding measure proved adequate for placement. Problems in separating true decoding errors from foreign language modification of English pronunciation may have reduced the value of the decoding test. In contrast, the failure of oral reading rate to predict placements is more puzzling. One possible explanation is that many of the non-native English speakers could comprehend English print in reading silently, but could not read aloud at a normal rate with acceptable accuracy.

Attempts to achieve extremely high accuracy in placement should be tempered by a consideration of the small number of placement levels usually available and by the lack of serious consequences for misplacement. Although moving students from one class to another after instruction has begun could be upsetting to some of the students involved, it is nevertheless an option, especially in consideration of the costs to students and staff of more lengthy testing. Teachers should be able to determine within a few class periods whether or not any of their students are seriously misplaced. Therefore, the penalty for placement failure may be relatively small. Furthermore, the reliability of the TABE Tests, particularly for administration upon entry to a program, is not so high that some students would not be misplaced, even with almost 3 hours of testing.

The data presented here supports the conclusion that the 37 minutes required by the TABE Locator Test will yield placement

22

results as reliable as the almost 3 hours of testing required by the Locator plus the Reading and Mathematics Tests. Furthermore, the Vocabulary Locator, requiring only 15 minutes, may also be a reasonable option for placement in many programs. The possibility that some test not evaluated here might perform better for placement than any of the measures actually evaluated also cannot be ruled out.[3]

Having scaled down the time required for placement testing to a low level, should testing be eliminated altogether for this function? Possibly, but this goes beyond the scope of the present study. Nevertheless, the possibility of self-placement can be raised, wherein examples of the materials used at each instructional level are provided so that students can decide on their own where they might best begin. The results of self-evaluation surveys, administered after each test in this study (see Venezky, Bristow, & Sabatini, 1993), offer some support for this idea, particularly for open-response tests like the TALS. Although a few studies have been done on self-evaluation of reading ability (Cross, 1981; Kaminsky & Hrach, 1990), much work remains to be done on the metacognitive abilities of adult literacy students before the limits of this approach can be determined (see Paris & Parecki, 1993). A small amount of instruction might allow most students to make reasonably reliable judgments of their abilities and at the same time demonstrate to them that they will be treated like partners in a learning process rather than like school children.

Overall, it may be concluded that less placement testing may be more valuable to both students and adult literacy programs. Less time on testing means more time on instruction. Perhaps even more important is that many adult learners have complained about their distaste for and fear of standardized tests. By cutting back on testing and moving toward a self-testing model, programs may stimulate greater motivation and satisfaction among the clients they wish to serve.

23

ENDNOTES

1 Exactly what would qualify as an external validation for class placement is not obvious. Teacher judgment is one candidate, but given the high degree of individualization that occurs in ABE and GED classes and the lack of sharply defined boundaries between them, teachers might be willing to accept students whose prerequisite abilities would place them at a lower or higher level. Further testing of prerequisites and class objectives might have been done, but within the bounds of general literacy objectives, teachers tend to adjust their instructional goals according to the abilities of their students.

2 For the TABE, which is a multiple-choice test, scores below what would be obtained by an average guessing strategy were not included in determining the four equal divisions.

3 For descriptions of other available adult literacy tests, see CASAS (1989) and Jackson (1990).
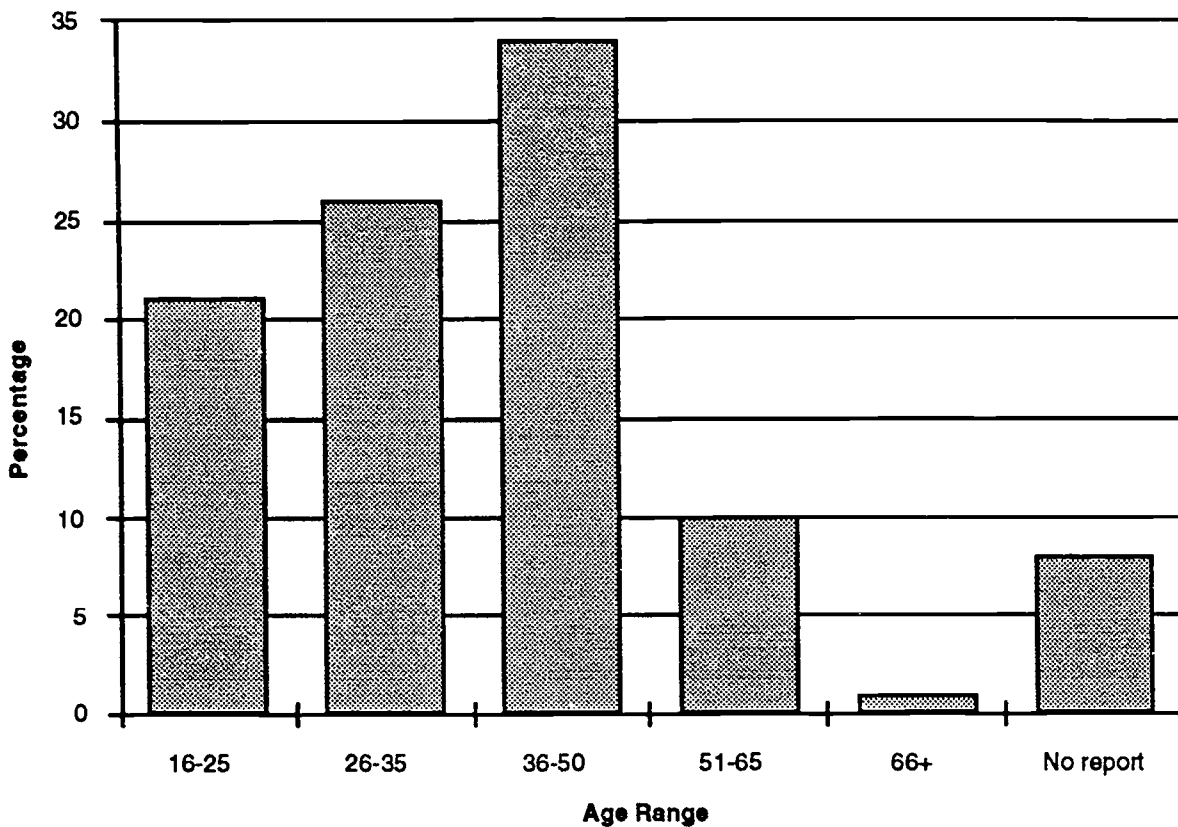
24

REFERENCES

Bristow, P. S., & Leslie, L. (1988). Indicators of reading difficulty: Discrimination between instructional- and frustration-range performance of functionally illiterate adults. *Reading Research Quarterly, 23*(2), 200-218.

Carver, R. P. (1990). *Reading rate: A review of research and theory.* San Diego, CA: Academic Press.

CASAS. (1989). *CASAS Adult Life Skills.* San Diego, CA: Author.

Cohen, J. (1960). A coefficient for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

Cross, K. P. (1981). *Adults as learners.* San Francisco: Jossey-Bass.

CTB/McGraw-Hill. (1987a). *Technical Report: Tests of Adult Basic Education (TABE) Forms 5 and 6.* Monterey, CA: Author.

CTB/McGraw-Hill. (1987b). *Tests of Adult Basic Education (TABE).* Monterey, CA: Author.

CTB/McGraw-Hill. (1987c). *Examiner's Manual: Tests of Adult Basic Education (TABE) Forms 5 and 6.* Monterey, CA: Author.

Development Associates. (1992). *Executive summary of the first interim report from the national evaluation of adult education programs.* Arlington, VA: Author.

Ehringhaus, C. (1991). Teachers' perceptions of testing in adult basic education. *Adult Basic Education, 1*(3), 138-154.

Jackson, G. B. (1990). *Measures for adult literacy programs.* Washington, DC: ERIC Clearinghouse on Tests, Measurement and Evaluation.

Kaminsky, S., & Hrach, E. (1990). The development of self-evaluation skills for improving reading performance. *Adult Literacy and Basic Education, 14*(1), 54-61.

Kirsch, I. S., & Jungeblut, A. (1986). *Literacy: Profiles of America's young adults.* (Report No. 16-PL-01). Princeton, NJ: National Assessment of Educational Progress.

Kirsch, I. S., Jungeblut, A., & Campbell, A. (1991). *ETS Tests of Applied Literacy Skills: Administration and scoring manual.* New York: Simon & Schuster.

Kirsch, I. S., Jungeblut, A., & Campbell, A. (1992). *Beyond the school doors: The literacy needs of job seekers served by the U.S. Department of Labor.* Princeton, NJ: Educational Testing Service.

Lytle, S. L., & Wolfe, M. (1989). *Adult literacy education: Program evaluation and learner assessment.* Columbus, Ohio: ERIC Clearinghouse on Adult, Career, and Vocational Education, Center on Education and Training for Employment, Ohio State University .

25

Nitko, A. J. (1989). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 447-474). New York: Macmillan.

Nunnally, J. C. (1967). *Psychometric theory.* New York: McGraw-Hill.

Paris, S., & Parecki, A. (1993). *Metacognitive Aspects of Adult Literacy* (Report No. TR93-9). Philadelphia, PA: National Center on Adult Literacy, University of Pennsylvania.

Simon & Schuster. (1990). *ETS Tests of Applied Literacy Skills.* Englewood Cliffs, NJ: Author.

Sticht, T. G. (1990). *Testing and assessment in Adult Basic Education and English-as-a-Second-Language programs.* San Diego, CA: Applied Behavioral & Cognitive Sciences, Inc.

Venezky. R., Bristow, P., & Sabatini, J. (in press). *Measuring gain in adult literacy programs: Enduring issues and a few answers.* (Technical Report). Philadelphia: National Center on Adult Literacy, University of Pennsylvania.

Willingham, W. W. (1974). *College placement and exemption.* New York: College Entrance Examination Board.

Whitney, D. A. (1989). Educational admissions and placement. In R. L. Linn (Ed.), *Educational measurement.* (3rd ed.) (pp. 515-525). New York: Macmillan.
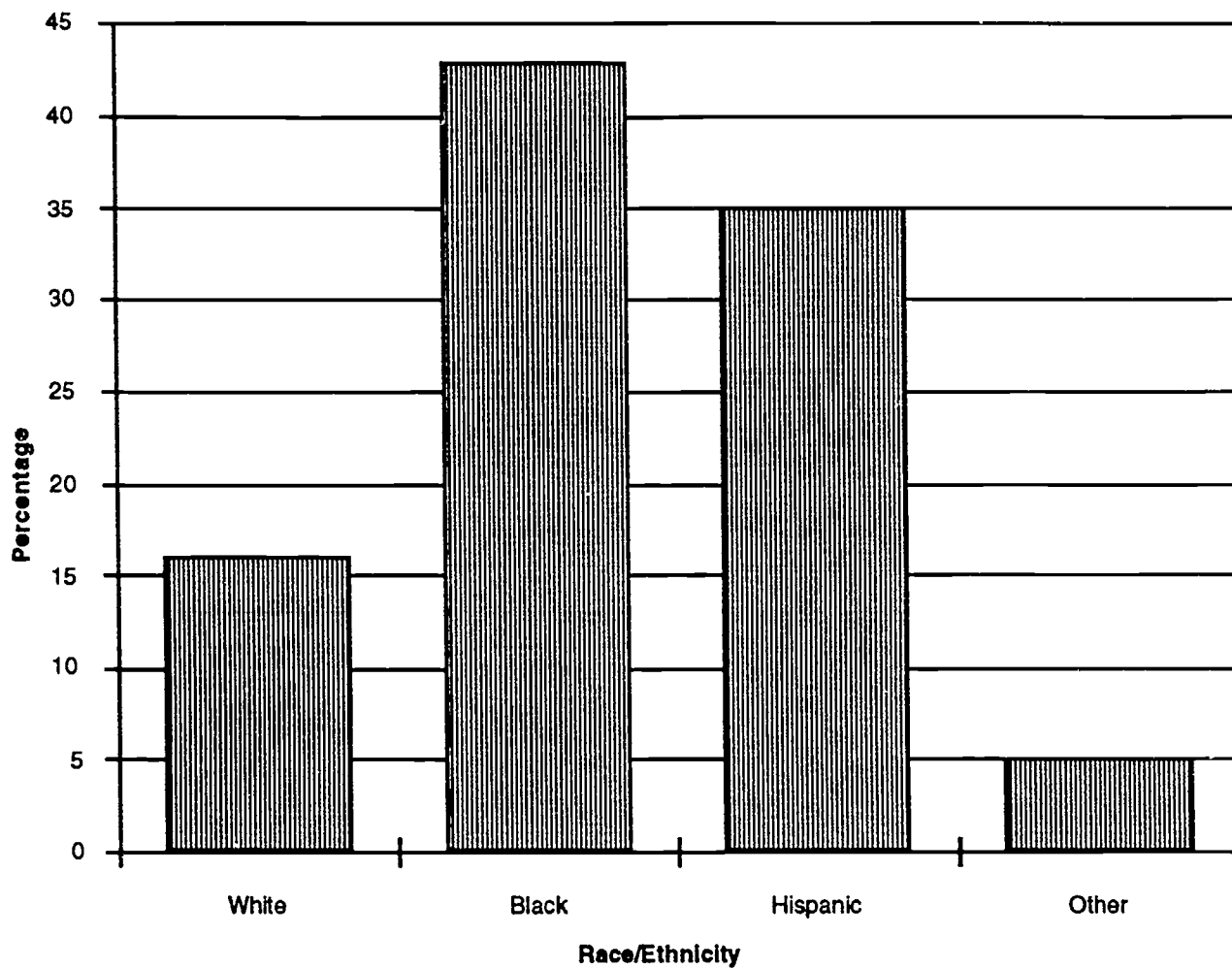
26

APPENDIX A

27

Figure 1: Age Distribution of Placement Students (N=168)

Figure 2: Race/Ethnicity of Placement Students (N=168)

29

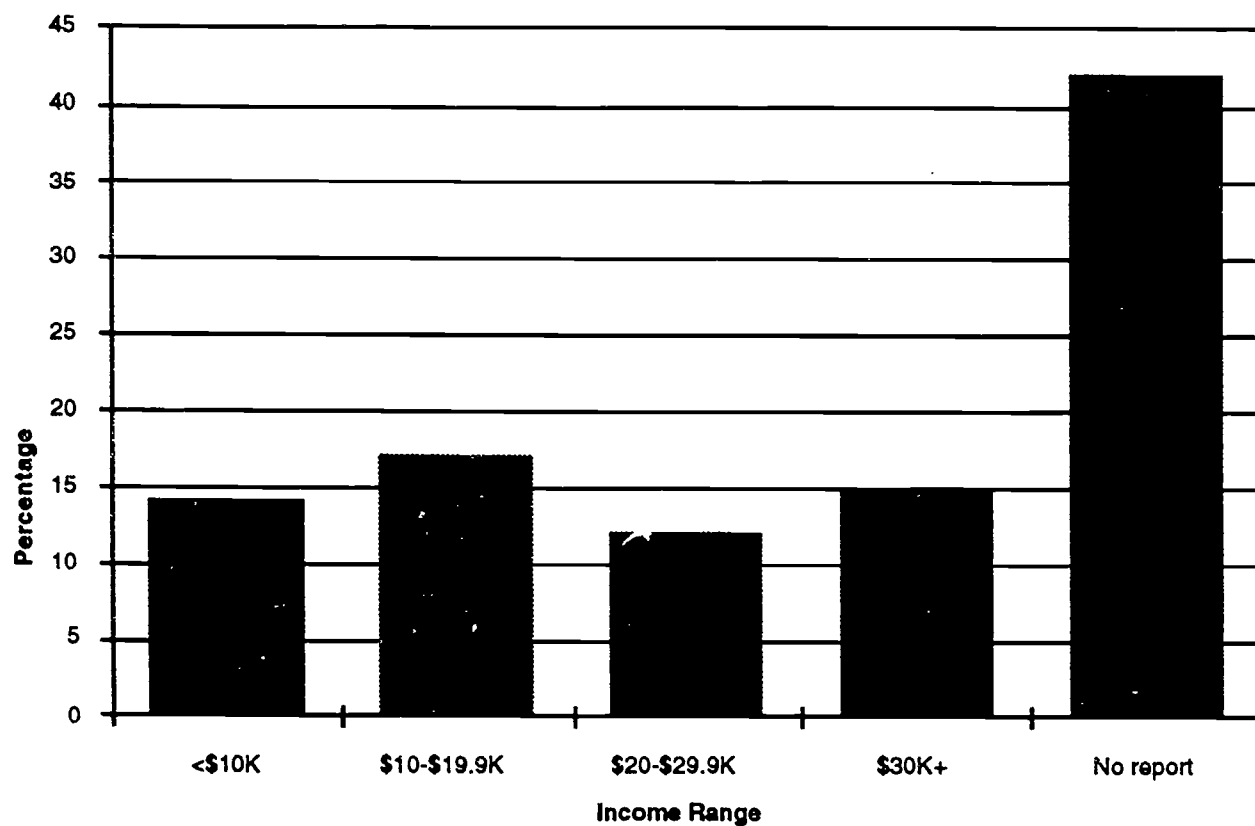**Figure 3: Income Distribution of Placement Students (N=168)**

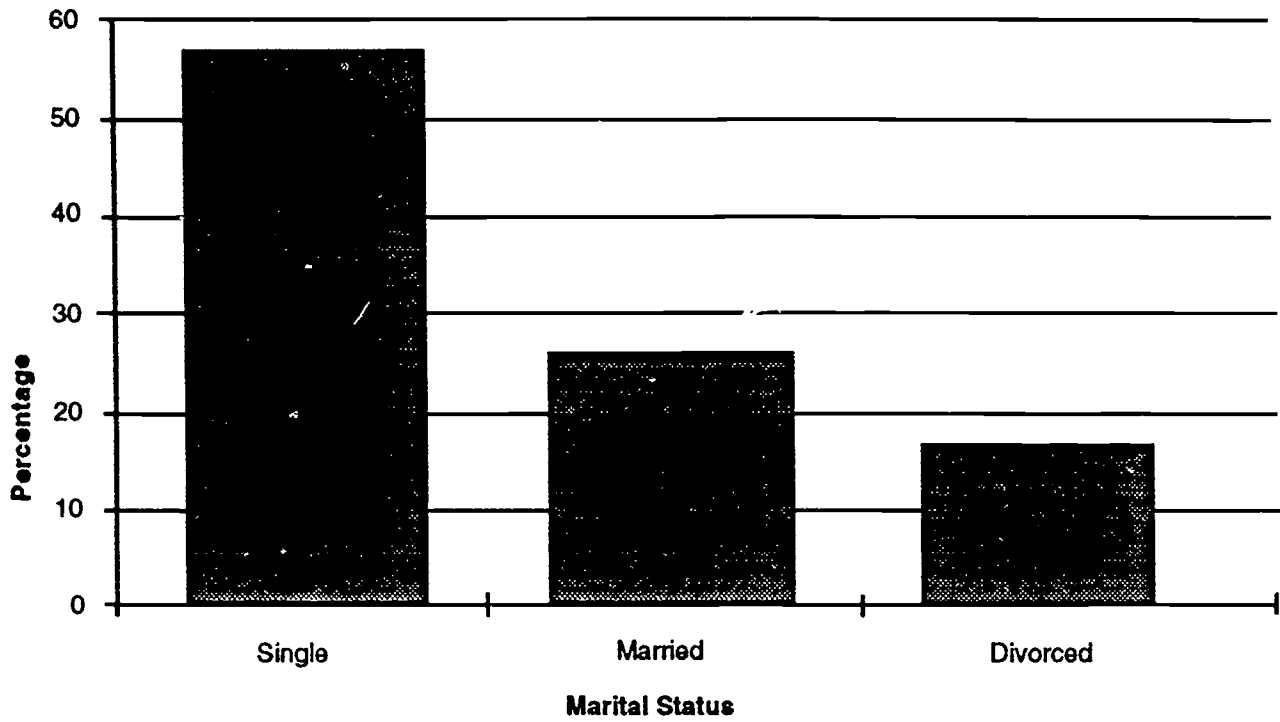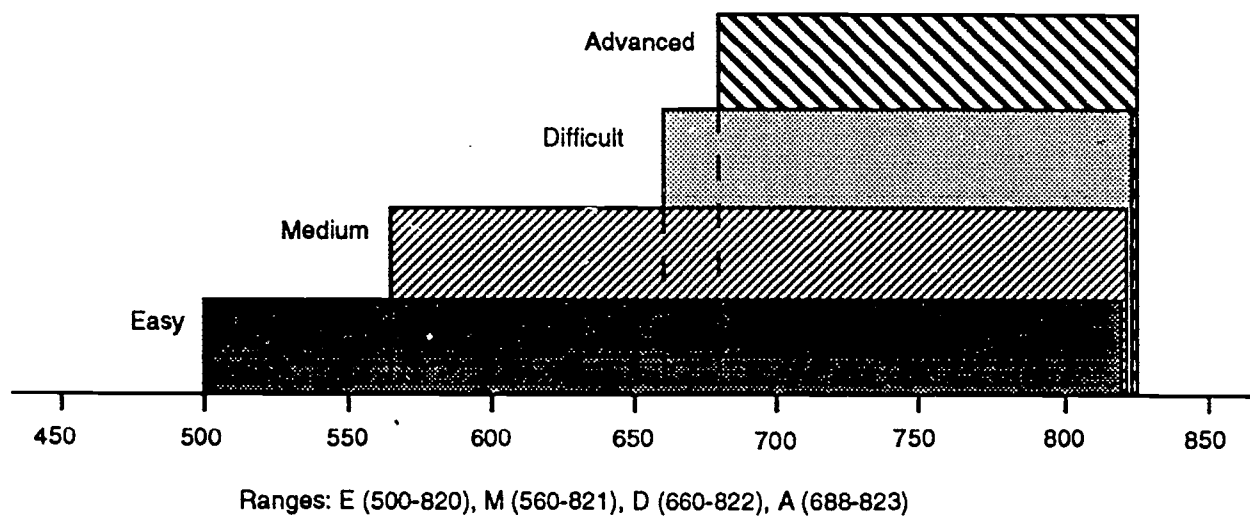Figure 4: Marital Status of Placement Students (N=168)

31

Figure 5: TABE Reading Comprehension Scale Ranges (Form 5)

Ranges: E (500-820), M (560-821), D (660-822), A (688-823)

32

APPENDIX B

Table 1: White Plains Placement Procedures, Fall 1991: Continuing Students

Continuing Students' Placement Levels Based on TABE Total Reading Results	Scores on Test Level Taken*				
	None	E	M	D	A
ABE 1	0 - 12 TABE Locator Total Score				
ABE 1		0 - 4.1	0 - 4.1	0 - 4.1	0 - 4.1
Take M Level*		4.2+			
ABE 2			4.2 - 5.9	4.2 - 5.9	4.2 - 5.9
Take D Level*			6.0+		
ABE 3				6.0 - 7.9	6.0 - 7.9
GED				8.0+	8.0+

* Students were placed based on their score on highest level taken.

CONTINUING STUDENTS PLACEMENT DESCRIPTION

For students who were continuing in the program, the normal placement procedures were followed. First, the TABE Locator Test was given. The total Locator score was used to determine which TABE Battery level would be administered initially (E, M, D or A) using the TABE authors' criteria. Also, students whose total scores on the TABE Locator Test were 12 or below did not take the TABE and were placed in ABE 1. The remainder of the students' class placements were based on their TABE Total Reading score computed from the TABE "Number-Correct to Subtest and Total Area Grade Equivalent, Complete Battery, Form 5, Reading" tables for each level. These tables combined students' Comprehension and Vocabulary Test scores to yield a Total Reading score in grade equivalents (GEs).

Students who took the E test were placed in ABE 1 if their GE scores were between 0 and 4.1. If their E Level scores were 4.2+ they were then given level M. All students who scored between 0 and 4.1 on the M test were placed in ABE 1; those who scored between 4.2 and 5.9 were placed in ABE 2. Students who scored 6.0+ on the M level test were given the D level test. Students who scored between 0 and 4.1 on the D test were placed in ABE 1. All students who scored between 4.2 and 5.9 on the D test were placed in ABE 2, and those who scored between 6.0 and 7.9 were placed in ABE 3. Students who scored at 8.0 or higher on the D Level were placed in GED. Students whose Locator Test scores indicated that they should take the A test were placed using these criteria: any students who scored unusually low (between 0 and 5.9 GE on level A) were placed in ABE 2; students scoring between 6.0 and 7.9 GE were placed in ABE 3; and those who scored at 8.0 or higher were placed in GED.

34

Table 2: White Plains Placement Procedures, Fall 1991: New Students

New Students' Placement Levels Based on TABE Total Reading or Comprehension Results	Scores on Test Level Taken		
	None	E	D
ABE 1	0 - 12 TABE Locator Total Score		
ABE 1		0 - 4.1	0 - 4.1
ABE 2		4.2+	4.2 - 5.9
ABE 3			6.0 - 7.9
GED			8.0+

NEW STUDENTS' PLACEMENT DESCRIPTION

For students who were new to the program and included in the study, modified placement procedures were followed. The TABE Locator total scores were used to determine which TABE Battery level would be administered using the TABE authors' criteria; however, only the E and D levels of the TABE were administered (as the design parameters of the study required). Students whose total scores on the Locator were 12 or below did not take the TABE and were placed in ABE 1. Students whose Locator scores indicated they should take the E or D levels were tested accordingly. Students whose Locator scores indicated they should take level M were randomly assigned to either the E or D level. Students whose Locator scores indicated they should take the A level took D.

Students who took the E level were placed in ABE 1 if their comprehension Grade Equivalent (GE) scores were between 0 and 4.1. Those whose comprehension GE scores were 4.2 or above were placed in ABE 2. The remainder of the students' class placements were based on their TABE Total Reading score computed from the TABE "Number-Correct to Subtest and Total Area Grade Equivalent, Complete Battery, Form 5, Reading" tables for each level. These tables combined students' Comprehension and Vocabulary Test scores to yield a Total Reading score in grade equivalents (GEs). Those whose Total Reading GE scores on the D test were between 0 and 4.1 were placed in ABE 1; those whose GEs were between 4.2 and 5.9 were placed in ABE 2. Those who scored between 6.0 and 7.9 GE on level D were placed in ABE 3. Students who scored at 8.0 or higher on the D level were placed in GED.

35

APPENDIX C

Table 1. Summary of Tests Administered for Evaluating Placement

Test Administered	# Items	Time Limits*
TABE Tests		
Total Locator	50	37
Vocabulary	30	19
Comprehension	40	8
Total Reading	70	57
Mathematics Computation	48	45
Mathematics Concepts & Appl.	40	43
Total Mathematics	88	88
TALS Tests		
Document Literacy	26	40
Quantitative Literacy	23	40
Prose Literacy	24	40
Oral Reading Tasks		
Decoding	42	10**
Passages	4	15**

* Time limits do not include preliminary activities such as instructions, distribution of testing materials, and practice items.
** Since no time limits were set on oral reading tasks, figures given are estimates.

Table 2. Decoding Test

List	Contents	Example
1	Simple CVCs	vun
2	CVC with consonant clusters	hent
3	CVC & CVCe with digraph consonants	shafe
4	CVC with digraph vowels	spawk
5	1 & 2 syllables with common prefixes and endings	refarbed
6	2 & 3 syllables with common prefixes and endings	impentive
7	4 & 5 syllables with common prefixes and endings	disfactible

(All lists have six items.)

37

Table 3. Means and Standard Deviations for Placement Mismatches

Test	n	Mean Mismatch	Standard Deviation
TABE			
Locator Total[1]	165	0.545	0.566
Total Reading[2]	168	0.613	0.661
Vocabulary	168	0.625	0.663
Comprehension	168	0.661	0.714
Locator Mathematics	165	0.661	0.743
Locator Vocabulary	165	0.667	0.682
Math Concepts and Applications	168	0.827	0.675
Mathematics Computation	168	0.851	0.845
TALS			
Document Literacy	168	0.839	0.710
Quantitative Literacy	168	0.893	0.724
Oral Reading			
Decoding	150	0.893	0.743
Rate	150	1.007	0.901

[1] Locator Total is the summation of the Locator Mathematics and the Locator Vocabulary raw scores.

[2] Total Reading is the average of the Reading Comprehension and the Vocabulary scale scores.

38

Table 4. Cohen's Kappa Coefficients for Placement Comparisons

Test	Kappa Coef.
TABE	
Locator Total[1]	0.32
Total Reading[2]	0.32
Comprehension	0.30
Locator Mathematics	0.30
Vocabulary	0.29
Locator Vocabulary	0.26
Mathematics Computation	0.17
Mathematics Concepts and Applications	0.07
TALS	
Document Literacy	0.12
Quantitative Literacy	0.07
Oral Reading	
Rate	0.18
Decoding	0.13

[1] Locator Total is the summation of the Locator Mathematics and the Locator Vocabulary raw scores.

[2] Total Reading is the average of the Reading Comprehension and the Vocabulary scale scores.

39

Table 5. Predictors of ABE/GED Placement: Initial Regression Model

Variable	Statistical Measures				
	Coef.	Std. Error	Std. Coef.	T	P(2 Tail)
Constant	-1.72	1.44	0.00	1.19	0.24
TALS Document Literacy	0.00	0.00	0.03	0.30	0.76
TALS Quantitative Literacy	0.00	0.00	-0.08	0.78	0.44
TABE Locator Vocabulary	0.06	0.02	0.27	3.28	0.00
TABE Locator Mathematics	0.10	0.02	0.54	4.80	0.00
TABE Vocabulary	0.00	0.00	0.23	2.41	0.02
TABE Comprehension	0.01	0.00	0.29	2.70	0.01
TABE Mathematics Comprehension	0.00	0.00	-0.20	1.81	0.07
TABE Mathematics Concepts & Applications	0.00	0.00	-0.13	1.19	0.24
Decoding	-0.02	0.0:	-0.10	1.58	0.12
Oral Reading Rate	0.00	0.00	0.10	1.27	0.20

Analysis of Variance for the Initial Model

Source	Variance Measures				
	Sum-of-squares	DF	Mean-Square	F-Ratio	P
Regression	151.42	8	18.93	40.23	0.00
Residual	73.39	156	0.47		

Dependent Variable: Four levels (ABE 1, ABE 2, ABE 3 & GED) treated as continuous variable.

N: 147*

Multiple R: 0.812; Squared Multiple R: 0.659

Adjusted Squared Multiple R: 0.634; Standard Error of Estimate: 0.69

* 21 of the 168 initial subjects were missing oral reading or decoding scores. They were eliminated from this model. In subsequent reduced models without oral reading and decoding scores, the largest initial subject set was used.

4()

Table 6. Predictors of ABE/GED Placement: Reduced Regression Model One

Variable	Statistical Measures				
	Std. Coef.	Std. Error	Coef.	T	P(2Tail)
Constant	-2.11	1.12	0.00	1.88	0.06
TABE Locator Vocabulary.	0.06	0.02	0.29	3.93	0.00
TABE Locator Mathematics	0.09	0.02	0.45	5.17	0.00
TABE Vocabulary	0.00	0.00	0.21	2.37	0.02
TABE Comprehension	0.00	0.00	0.25	2.94	0.00
TABE Mathematics Computation	0.00	0.00	0.00	2.59	0.01

Analysis of Variance for the Reduced Model

Source	Variance Measures				
	Sum-of-squares	DF	Mean-Square	F-Ratio	P
Regression	150.04	5	30.01	63.81	0.00
Residual	74.77	159	0.47		

Dependent Variable: Four levels (ABE 1, ABE 2, ABE 3 & GED) treated as continuous variable.

N: 165*

Multiple R: 0.817

Squared Multiple R: 0.667

Adjusted Squared Multiple R: 0.657

Standard Error of Estimate: 0.69

*3 of the 168 initial subjects were missing Locator scores. They were eliminated from this model.

41