DOCUMENT RESUME

ED 360 803 EC 302 385

AUTHOR MacMillan, Donald L.; And Others
TITLE A Study of Minimum Competency Tests and Their Impact. Final Report.
INSTITUTION California Univ., Riverside.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
PUB DATE [90]
CONTRACT G008530208
NOTE 162p.; Appendices B and C may not reproduce clearly due to filled print.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Academic Achievement; *Academic Failure; Attendance; Discipline; Dropout Rate; *Educational Attainment; *Educational Experience; *High Risk Students; Junior High Schools; Junior High School Students; Low Achievement; *Mild Mental Retardation; *Minimum Competency Testing; Self Concept; Student Attitudes

ABSTRACT

This project studied the impact of Minimum Competency Tests (MCT) on the educational careers of secondary students. The educational experiences of those students passing the MCT were compared with the experiences of those students not passing the MCT. The study compared three categories of students (learning handicapped, educationally marginal, and controls) on such outcomes as types of courses taken, transferring out of the district (to a district with an "easier" test), early school dropout, school attendance, disciplinary problems, self-concept, and attitudes toward school. Results indicated that performance on the MCT greatly impacts the types of courses selected, with students passing the MCT enrolling in more elective courses and fewer remedial courses. Students failing the MCT transferred up to 10 times more frequently than those passing the MCT. Students failing the MCT were up to 10 times more likely to drop out of school than passing students. Students failing the MCT were absent more often and had lower self-concept in the areas of academics in general, math, honesty, and same-sex relations. Appendixes include a sample data collection sheet, a self-description questionnaire, and a school attitude survey. References accompany each chapter. (JDD)

A Study of Minimum Competency Tests and Their Impact

FINAL REPORT

Grant #G008530208
Office of Special Education and Rehabilitaion Services
U.S. Department of Education

Donald L. MacMillan
Irving H. Balow
Keith F. Widaman
Rita E. Hemsley

University of California, Riverside

A Study of Minimum Competency Tests and Their Impact

FINAL REPORT

Grant #G008530208
Office of Special Education and Rehabilitaion Services
U.S. Department of Education

Donald L. MacMillan
Irving H. Balow
Keith F. Widaman
Rita E. Hemsley

University of California, Riverside

Acknowledgments

The authors of this report are indebted to a number of individuals who assisted in varied ways during the conduct of the research reported herein. Project findings reflect the efforts and cooperation of many individuals whose contributions the authors wish to acknowledge. The students in each of the cooperating districts patiently endured our annual assessments and we want to thank them for their cooperation. Furthermore, we wish to thank their parents/guardians for granting the necessary consents as a reflection of their support or faith in the significance of the research.

Since the inception of the project, we have tried to protect the identity of the participating students as well as that of their schools and districts. In keeping, we must refrain from acknowledging by name the districts and individuals in those districts who were instrumental to the conduct of this research. Nevertheless, we do recognize the efforts of a great number of individuals: district level administrators and staff; administrators, counselors, staff and teachers at school building sites. Without their assistance and cooperation the project could not have been completed. The project was very labor intensive and the intrusion into offices, classrooms, and records required project staff to be in almost constant contact with personnel in the cooperating districts. We sincerely extend our appreciation to them for their willingness to assist and for their pleasant demeanor throughout a lengthy period of years.

Over the course of the five year project, we employed a number of students at the University of California, Riverside. During summers we also were fortunate to secure the services of teachers in the local schools to assist in the recording of information from school records. Without exception, the quality of their work and the way in which they conducted themselves was exemplary. While their responsibilities varied (data collectors of information from school records, data entry, testing groups of students), the following individuals participated in some time and in some way during the project: Heather Ariail, Jason Berry, Sherry Best, Selena Blaylock, Peter Buss, Rachel Duclos, Sheila Eiden, Cathy Espinoza, Andie Farren, Michael Fitzwater, Karen Fleck, Kerry Flynn, Trish Fuller, Susan Haimowitz, Patty Hawley, Pat Hieda, Janet Leach, Lanny Little, Mary Kay Moersch, Terry O'Beirne, Nancy Sidor, Tara Sneed, Ann Spencer, Marsha Thicksten, and Erin Turner. If we have omitted any names, we sincerely apologize.

Throughout the project, Peggy Darlington served in various critical roles: assuming all secretarial duties, managing budgets, coordinating preparation and duplication of instruments, assisting in data collection and entry as needed, and being on the "other end of the phone" when school personnel called with questions or problems. During the first three years of the project, the expertise of Dr. Todd Little was frequently sought for problems in analyses and data processing. He assisted our staff and saved us in our dealings with a huge data set through his design of the data system as well as with his trouble-shooting preparations for computer runs. We would also like to acknowledge the role of Karen Fitzhugh in the early phases of the project. Karen was one of the first project coordinators who assisted in

maintaining contact with the schools, scheduling, supervising data collectors, and otherwise insuring that what needed to be done was, in fact, done. Finally, Dio Saucedo's help was sought at various times. He assisted in translating written communications for Spanish-speaking families in our study, made phone calls to Spanish-speaking families as needed, and assisted in data collection and testing at various times.

We would also thank the support provided by the administration of the School of Education at the University of California, Riverside.

To all of the individuals listed above, we express our gratitude. Their contributions improved the quality of this report. Any shortcomings contained in this report are assumed by the authors alone.

TABLE OF CONTENTS

CHAPTER III
Psychometric Characteristics of the Tests

CHAPTER IV
Consequences of MCT Performance on Student Careers

LIST OF TABLES

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

LIST OF APPENDICES

CHAPTER I

Introduction and Background

The significance of this research derives from the collision of two independently enacted educational policies. The impetus behind both policies is heavily weighted with sociopolitical factors. While both were clearly intended to benefit school children, it is our belief that when considered together they harm, and will continue to harm, a substantial segment of our public school population; namely learning handicapped (LH) and the educationally marginal students.

The two policies of concern are the changes in identification of mildly handicapped learners, resulting in more restrictive definitions (particularly when applied to ethnic minority children), and mandated minimum competency tests (MCT), which must be passed in order to graduate from high school. The former policy can be characterized as a push for equity in educational opportunity while the MCT policies appear to push for excellence. Caught between these policies are public school students who are mildly handicapped (educable mentally retarded, EMR; learning disabled, LD) and another segment of low achieving "normal" students whose handicaps, if they exist, go undetected or ignored (e.g., when the student in question is a minority student and the district is "at quota" for his/her particular minority group).

Olsen (1980) observed that much of the opposition to MCTs may be traced to the fact that ". . . in the scramble to tool up minimum competency testing programs to respond to societal pressure handicapped and minority populations were seldom considered" (p. 177). Society became outraged at increasing school costs and declining test scores. Some lost faith in the high school diploma being a certificate of competence. Many others shared a concern over

deteriorating educational standards and demanded that steps be taken to guarantee reasonable standards, if not excellence. One step in that direction is the MCT movement (Haney & Madaus, 1978), as most states have some sort of MCT as a determinant of which students will receive high school diplomas and, in some cases, which students are advanced from grade to grade. Such policies have been enacted despite heated exchanges over the feasibility of developing reliable and valid tests for these purposes (see Anderson & Lesser, 1978; Brickell, 1978; Fisher, 1978; Glass, 1978; Haney & Madaus, 1978; Hart, 1978; Wise, 1978).

Overlapping temporally with the push for excellence has been the push for equity in educational opportunity for minority and handicapped children. In the case of mildly handicapped children (particularly EMR), the issue of equity for the handicapped was superimposed on the issue of equity of educational opportunity for ethnic minority children, since disproportionately high number of EMR students were minority children. Court cases such as Diana v. State Board of Education (1970) and Larry P. v. Wilson Riles (1971), both tried in California, exemplify plaintiffs seeking equity for minority children classified as EMR. Many of the points raised by plaintiffs in these cases were incorporated into PL 94-142 (1975), the federal legislation concerning the education of handicapped school children.

The net effect of the court actions and legislation has been a drastic reduction in the number of EMR students. For example, in California from the time the Diana litigation began (July, 1969) until October of 1977 the total number of EMR children declined from 55,519 to 19,370 (Lambert, 1981). This downward spiral in the number of EMR has continued. Forness (1985) showed a 30% decline in California school children classified as mentally retarded

(this is not broken down into EMR, TMR, etc.), from 42,900 to 29,900, between the academic years 1976-77 and 1981-82. These figures suggest that in California schools today many fewer children are served as mentally retarded and, conversely, a number of the "former EMR type" of student are not receiving any special education services and are enrolled as "normal" students in general education.

The press for equity regarding the mildly retarded has taken the form of avoiding false negatives -- that is, avoiding labeling children who are not "really retarded." Stated differently, if the system is to err, policy appears to favor not labeling some children in need of special education services as preferable to labeling some erroneously.

The presses for equity and excellence are on a collision course, particularly as they pertain to marginally achieving students and mildly handicapped students. For example, in their review of all efficacy studies and mainstreaming studies, Semmel, Gottlieb, and Robinson (1979) concluded that the reading achievement of mentally retarded subjects (sometimes with IQs up to 85) was exceedingly poor; and the wrote, "mean reading scores of EMR pupils never reached a grade level of 4.0" (p. 237). Yet, MCT levels for passing are more frequently set at grade equivalents of 7.0 or 9.0. The prospects for students like the former EMR, (IQ up to 85) and the current handicapped students to pass MCTs appear bleak.

Major decisions regarding students' careers are being made based on these tests of questionable reliability and validity. Those certified as handicapped can take these tests under "modified" conditions despite the fact that we do not know what these modifications do to the already shaky reliability and validity of the tests. Or, handicapped students, under

certain circumstances, can be declared "competent" and granted a diploma by meeting differential standards (often completing IEP objectives). For those marginal achievers who avoid labeling as handicapped, the MCT must be taken under standard conditions (i.e., modifications are not permitted) and they will be denied high school diplomas unless the test is passed.

General Problems in Using MCTs

"Minimal competency testing for high school graduation and grade-to-grade promotion continues to be one of the most explosive issues on the educational scene today" (Pipho, 1978, i). This concern summarizes concerns expressed by many regarding the MCT movement in general, and as it applies to students who are not considered handicapped in any sense of the word. MCTs have been constructed, administered, scored, and used to decide who will, and will not, be promoted or granted a diploma. Yet, some fundamental problems concerning the construction and use of MCTs have not been addressed satisfactorily.

Enthusiasm exists in some circles for the use of MCTs, largely fueled by increased public and legislative concern over the questionable quality of high school graduates. Various authors mention concerns held about public education, such as deteriorating educational standards reflected in declining SAT scores (Haney & Madaus, 1978), a dearth of data regarding school effectiveness (Fisher, 1978), the public outcry against declining student mastery of basic skills (Candor-Chandler, 1978), the need to make the public schools more accountable (Wise, 1978), the need to prevent passing incompetent students on the basis of social promotion (Haney & Madaus, 1978), and the denial of needed instruction for students who are incompetent (Hart, 1978). Concerns such as these resulted in rapid-fire passing of MCT legislation at

the state level. Cawelti (1978) reported that at the beginning of 1977 only seven states had passed competency based laws, but by the end of 1977 this total had risen to 30 states and all others were considering similar legislation.

Haney and Madaus (1978) observed that the focus on testing was primarily instigated by noneducators as a means of making public education accountable and to assure taxpayers that they would receive a "return for the dollar." However, the tests that resulted have been assailed by scholars in a variety of ways. Among the questions raised are: What do they seek to measure? How do they measure competence? How high are the minimums set? When in a student's career should competencies be measured? What is to done with the incompetent?

What is measured by the MCT?

The test themselves can vary in terms of whether they tap basic skills, school subjects, life skill areas, basic skills applied in school subjects, or basic skills applied in life skill areas (Brickell, 1978). Furthermore, once this decision is made, the issue is further complicated by the fact that school success is not a good predictor of success in later life. The attractive product of a MCT is an education that develops competencies needed for life. In reality, this amounts to a "bait and switch" tactic. The product that one assumes is being offered is a life skills curriculum that will develop the competencies needed for survival; the product that one receives is a score on a paper and pencil test that is a flimsy pretense for the real thing: accomplishment of the tasks in day to day living.

The decision regarding what specific competencies should be measured and how they should be measured is an imposing problem, and one which has not been

resolved satisfactorily.

Measuring competence. The most common way competencies are measured is with paper and pencil tests which are easier and cheaper to administer, however they possess very limited predictive validity for adult success. As Nathan and Jennings (1978) observed, the most important qualities of life do not get answered in paper and pencil tests, which fail to address the issue of the relevance of the school curriculum.

Another telling critique of extant paper and pencil tests used as MCTs was offered by Glass (1978). In his review of the Florida test, developed by Educational Testing Service (ETS), Glass noted that in examining a half dozen methods for establishing criteria (i.e., mastery levels, standards, cut-off scores) for passing, he found them all to "yield arbitrary and dangerous results." Furthermore, as he examined the item development, he was stunned to find that ETS failed to produce items to any statistical prespecifications -- i.e., with a particular mean and variance. In summary, the items and the criteria for passing are arbitrarily determined.

Setting minimums. The performance level for passing is arbitrarily set for declaring a student as possessing the minimal competencies, or as Haney and Madaus (1978) stated, there is simply no scientific basis for determining cut-offs. Rather, these decisions are inevitably politically based. These same authors related reactions to arbitrarily set cut-offs. For example, when the New York State Board of Regents proposed a ninth-grade equivalent achievement, one official in the mayor's office responded, "What happened to the twelfth grade?" while a reaction in Connecticut to a seventh-grade equivalent reading level was "We're paying for twelve years of schooling, but we're only getting seven years" (Haney & Madaus, 1978, p. 468). While these

reactions reflect a misunderstanding of grade-equivalents, the fact that one test utilizes 9th grade and one 7th clearly evidences differences in opinion regarding what constitutes minimum competence. In Florida the results yielded 8% failing the communication test and 36% (high school juniors) failed the math test. While the inference most widely made for these results was that math skills were relatively poorer than communication skills, Glass (1978) suggested the possibility that the math items were on average more difficult than the items on the communication test. Hence, not only might overall grade equivalents required for passing vary from test to test, they many vary from subtest to subtest within a test.

Timing and remediation. In the proposed research, MCTs are being studied as they are used to determine who will be granted a diploma --i.e., who will graduate from high school. However, test results are used in a diagnostic-prescriptive fashion as they are initially administered to 8th grade students and those who fail are programmed into remedial instructional classes. Haney and Madaus (1978) caution that this could influence what is taught -- i.e., teachers might teach to the test. These same authors argue that the issue of greater relevance is what should be learned that is not being learned presently, and what should be taught which is not being taught presently.

For students who fail MCTs, there are remedial programs designed to assist them in almost all states (Gorth & Perkins, 1979; Lewis, 1979). These remedial programs are bound to increase the expense already resulting from implementing MCT provisions. Andersen and Lesser (1978) provided some cost figures: (a) Washington, $43 million and $47 million for remedial programs in math and reading, respectively, (b) New Jersey, $70 million for compensatory

programs in 1978-79, (c) Michigan, $28 million, and (d) Florida, $10 million.

Minimal Competency Testing and the Handicapped

The issue of how, if at all, MCT should be utilized with handicapped children is far from resolved (see McKinney, 1983; Olsen, 1980; McCarthy, 1980; NASDSE, 1979; Ross & Weintraub, 1980; Safer, 1980). Positions advocated include standard administration of the tests (after all they are minimal competencies), modifying the procedures to accommodate the specific disability, waiving the MCT and accepting the completion of the child's program of study (outlined in the filed Individualized Education Plan). No single position appears to have widespread support, yet reports from states such as North Carolina (McKinney, 1983) and Florida (Safer, 1980) suggest that modified test protocols are being used rather widely.

Since the handicapped students of interest in this study are mildly handicapped learners, those modifications specifically designed to accommodate children with sensory or physical disabilities are not considered. McKinney (1983) described several modifications which are appropriate for use with mildly handicapped children, including: (a) extended time limits, (b) permission to write directly in the test booklet, (c) recording of answers by the proctor, (d) small group administration, and (3) audio-cassette editions of the tests.

McCarthy (1980) suggested as an alternative to requiring the child to pass a MCT, acceptance of successful completion of the IEP goals as evidence of "competence." This approach could be seen as discriminatory against nonhandicapped students who complete their program of study successfully (i.e., pass required courses for graduation) but are denied a diploma when they fail to pass the MCT. Furthermore, McCarthy noted that should a

handicapped child exhibit "deficiencies" on a MCT, one might argue that they should be remediated lest they (i.e., the deficiencies) later prove a handicap to the person occupationally or socially. The compatibility of the IEP and MCT may be a function of the extent to which academic skills vs. life adjustment are stressed on each (Olsen, 1980). Furthermore, the values held by the team writing the IEP (i.e., the emphasis placed on reading and math vs. social and interpersonal skills) could lead to students being declared "competent" on skills (e.g. interpersonal skills) which would not permit them to pass the MCT (assuming it stressed academic tasks).

In California there is no uniformity in policy regarding how special education students are to demonstrate competence. Guidelines from several school districts attest to the fact that special education students can meet competency standards in a variety of ways available to no other students. What appears in these rules and regulations are alternative routes by which a special education student can be certified as "competent" and granted a high school diploma: (a) He/she can take, and pass, the MCT, (b) He/she can take the MCT under modified test conditions and pass the test, or (c) the district can waive the MCT and establish other (than the MCT) standards (sometimes referred to as Differential Standards) determined for the student by the IEP team. In cases (b) and (c) there remains the question of the equivalence of competence as determined by these options vis-a-vis passing the MCT under standard testing conditions.

Modified test procedures (i.e., "Alternate Modes") include: (a) testing by the special education teacher, (b) reading test directions to the student more than once, (c) extended time to take the test within the testing day, (d) giving breaks to the student during the testing time, (e) using paper markers

to isolate questions and answers, and (f) record answers in the test booklet (others are available for students with sensory and/or motor handicaps). As noted previously these modifications need to be studied in order to ascertain what effect they have on the reliability and validity of the tests. Clearly, as Olson (1980) noted, the MCT movement, in its enthusiasm, did not consider that the handicapped and the "adjustments" being made, are in need of study.

The use of Differential Standards by which to certify the handicapped as competent raises a host of questions, yet unresolved. McCarthy (1980) raised the charge of reverse discrimination. How can one certify a handicapped child as competent based on completion of IEP goals, and yet declare a regular class student incompetent because he/she fails to pass the MCT despite having completed successfully his/her program of study? This issue arises with our EM sample which may, in considerable numbers, "pass" courses yet fail to pass the MCT. For these students (EM) the alternatives available to handicapped students are not available. The most obvious question, however, is how the required passage of the MCT can be waived if, in fact, the test measures minimum competencies that any recipient of a diploma should demonstrate. McCarthy (1980) also noted that if deficiencies in the academic repertoire of mildly handicapped children are uncovered by the MCT, they should be remediated, rather than be ignored, lest they prove later be an impediment to the adaptation of the handicapped child as an adult.

Clearly, the responses to problems encountered as MCT has been extended to the handicapped have been stop-gap measures which raise as many questions as they answer. The need to investigate the impact of these modifications seems readily apparent.

Psychometric Issues with Mildly Handicapped and Educationally Marginal Students

Minimum competency tests (MCT) have been legislated as requirements in 40 states in an effort to improve instruction, and consequently achievement, in public schools. In many states, a student must pass the MCT to be awarded a high school diploma. The educational (e.g., Klein, 1984; Serow, 1984; Walstad, 1984) and legal (e.g., Logar, 1984; McCarthy, 1983; White, 1984) issues associated with MCTs have been discussed extensively in the literature.

Among the states requiring MCTs there is variability both in the way proficiency is assessed and in the performance standards required for passage. In some states (e.g., New York, North Carolina), a statewide MCT is used and a common standard applied, while in others (e.g., California) the tests and standards for passing are determined by local school districts (Jaeger, 1982). The rationale for local option testing includes the possibility that tests can be tailored to the curriculum of the local district and that any single statewide test violates the right of local school boards to establish criteria for graduation. In California, each school district is required to adopt or develop MCTs in reading, writing, and mathematics; establish standards for passage; and require passage before a student may graduate. In addition. the school district must provide remediation to students not making sufficient progress and reassess those students. In the next few pages we examine some psychometric issues arising in local option states, such as California, particularly as they affect mildly handicapped students (particularly the learning disabled, who represent the majority of this group) and those termed "educationally marginal". Let us first define what we mean by "educationally marginal".

Educationally Marginal Students. As a result of the change in definition of mental retardation (Grossman, 1973) that followed court cases challenging the diagnostic practices for identifying educable mentally retarded (EMR) students (e.g., Diana v. State Board of Education, 1970; Larry P. v. Wilson Riles, 1971), those students with IQs between approximately 70 and 85 were deemed no longer eligible for classification as EMR. Essentially, these decisions reflected the belief that those students were afforded greater equality of educational opportunity in regular programs than in special programs. It has been noted (Forness, 1985; Polloway & Smith, 1988), that children scoring in the IQ range of 70 to 85 are not eligible for services as learning disabled (LD) because a "severe discrepancy" between aptitude (i.e., IQ) and achievement cannot be demonstrated because, regardless of how low achievement falls, it is explained in terms of the low IQ. As a result, there is a group of students in the regular classes who psychometrically were eligible for EMR before the change in definition, but are now considered "normal", albeit educationally marginal. MacMillan, Hendrick, and Watkins (1988) have noted elsewhere that this group has been caught in the conflict between policies designed to achieve "equality" on the one hand (e.g., mainstreaming) and others designed to achieve "excellence" (e.g., minimum competency testing).

These educationally marginal children are afforded none of the protections afforded handicapped students when it comes to demonstrating proficiency or minimum competency. For example, handicapped students in California can be judged competent in one of three ways. First, they can pass the MCT under the standardized administration procedures and score above the standard for passing. Second, they can take the MCT under modified testing conditions that

accommodate their specific disability (e.g., no time limits, instructions read by examiner, answers recorded by the examiner). Finally a handicapped student can be judged by the IEP team to meet the requirement according to "differential standards." This typically translates into demonstrating mastery of the long-term objectives listed in the IEP for the student. The educationally marginal student, however, must pass the MCT under standard administration conditions. Unless this is accomplished, the student will be denied a high school diploma, which reflects and reinforces the perception of "normal" student.

As we consider the specific issues, we will relate them to these two groups of students, handicapped and educationally marginal.

MCTs in Local Option States

In states using a single statewide MCT, test construction has typically been quite sophisticated, involving experts in the test development, extensive field testing, item revision, and establishing standards for passing. As a result the MCTs finally used in those states are "known quantities" with psychometric properties that are respectable and defensible.

In local option states, however, there, is greater variability in the psychometric properties of the tests used. Each district is required to develop its own test or adopt an existing instrument. According to the California State Department of Education report, Proficiency Assessment in California (1980), 78% of the school districts developed their own tests and 46% of these wrote their own items or secured items from unknown sources.

Expertise in testing varies considerably from district to district. Some of the larger districts have well-developed departments of research and evaluation, while some smaller districts have no such formal departments;

responsibility for test development falls to someone with very limited expertise. In addition, some districts avail themselves of consultants from county offices or nearby universities, while others do not. As a consequence, the sophistication with which any two given MCTs were developed may vary greatly. Some districts simply adopted a standardized achievement test and adopted a "grade equivalent" score to constitute a standard for passing. Others gathered personnel and developed items reflecting what they thought everyone should know or be able to do, with almost total disregard for difficulty levels reflected in the various items. Yet others engaged in very sophisticated efforts to develop tests sensitive to the curriculum of the district with reasonable standards for passing. Let us consider the psychometric issues that arise in MCT development, particularly as they pertain to handicapped students and those of marginal ability. For these two groups, the psychometric problems are frequently compounded.

Reliability of MCT

Illustrative Data on Reliability for Seven Districts

One fundamental issue regarding MCTs in local option states concerns the reliability of the test being used to determine the short-term educational future of the students. Reliability refers to the consistency with which a set of test scores measure what they purport to measure -- in this case, proficiency. It is expressed in terms of a correlation coefficient reflecting stability, equivalence, or internal consistency (Mehrens & Lehmann, 1984). Reliability tends to increase as the number of test items increases.

The issues concerning reliability can be illustrated by data on MCTs derived from seven California school districts varying in size (about 3000 to over 50,000), location (Bay Area, Central Valley, Southern California),

setting (urban, suburban, rural), and growth characteristics (declining enrollments, steady, increasing enrollments). Each district provided results from standardized achievement tests and MCTs for the first year of administration. By using MCTs for the first year, the greatest variability in scores was achieved, which provided the highest estimates of reliability and clearest separation at the criterion level on the proficiency test.

Reliability was estimated from raw scores using the Kuder-Richardson Formula 21, which is a measure of homogeneity of test items -- the extent to which each item measures the same thing as the other items. Because the error term was limited to events occurring in the testing situation (i.e., no error associated with stability of response over time), a higher estimate of reliability is secured than would result from using the parallel forms method. Table 1.1 displays the KR21 reliability for each section of the proficiency test (reading, writing, and mathematics) and the corresponding sections of the nationally standardized test used in each district. In general, the reading subtests of the MCTs tend to be less reliable than the nationally standardized tests, while the writing and mathematics subtests tend to be as reliable as the nationally standardized tests used. Note the low MCT reliabilities for District A in all areas and District C in reading, neither of which resulted from restricted variability.

Table 1.1

MCT and Standardized Test KR_{21} Reliabilities for Seven School Districts in Three Areas

	Reading		Writing		Mathematics	
District	MCT	Stand.*	MCT	Stand.	MCT	Stand.
A	.70	.92	.63	.92	.66	.95
B	.86	.94	.92	.80	.91	.90
C	.58	.89	.86	.86	.90	.92
D	.93	.91	.89	.80	.92	.90
E	.90	.94	.93	.81	.92	.78
F	.94**	.98	.94**	.95	.92	.99
G	.87	.95	.94**	.94	.95**	.96

Note. The KR_{21} reliabilities shown are not those in the technical manuals, but are the reliabilities calculated on the data from the same groups of students used to calculate the MCT reliabilities.

* Nationally standardized achievement test used in the respective districts.

** Meets Kelly's standard for adequate reliability to evaluate level of individual accomplishment.

How Reliable Should MCTs Be?

Visual inspection of the reliability coefficients for MCT subtests suggests many are in the .80 to .90 range. How high should they be? The answer depends upon the decisions one wants to make from the test results. Long ago, Kelly (1927) devised a set of standards based on the assumption that a test must make discriminations of a difference small as .26 times (approximately one-fourth) the standard deviation of a grade group with a chance of 5 to 1 of

being correct. Following are the minimum reliabilities for four purposes: (a) .50, to evaluate the level of group accomplishment, (b) .90, to evaluate differences in level of group accomplishment in two or more performances, (c) .94, to evaluate level of individual accomplishment, and (d) .98, to evaluate differences in level of individual accomplishment on two or more performances. Mehrens and Lehmann (1984) advised that, "If the decisions the scores will help make are extremely important and/or irreversible, then the reliability of the measure is of more concern than if the decision is not quite so important and/or is tentative and reversible" (p. 285).

MCTs are used to make very important decisions, and usually irreversible ones, based on passing or failing. In many districts, failure results in a student being programmed into a remedial course for each subtest failed on a daily basis for at least one semester. Decisions about who will receive diplomas are thus very important decisions.

What would it take to improve the reliability of the MCT subtests to .94? The Spearman-Brown prophecy formula permits determination of how much longer a test would have to be to change from a measured reliability to a higher reliability by using equally good items. The formula is:

$$r_n = \frac{nr_s}{(n - 1)r_s + 1}$$

where n = number of times the length of the test will be multiplied

s = present length of the test

A test with a reliability of .89 would have to be twice as long to have reliability of.94; a test with a reliability of .80 would need to be four times as long, and a test of .70 would have to be eight times as long to reach a reliability of .94. Obviously, improving the quality of the items could

reduce the extent to which the length of the test would need to be increased in order to achieve this acceptable reliability (.94).

The decisions being made on the basis of MCTs are very significant, and the evidence on which these decisions are being made must be more reliable than appears to be the case at present.

Accuracy of Decisions

The reliability of a test determines the accuracy of decisions made based on test performance. The MCT reading subtest used in District C has an estimated reliability of .58, with a mean score of 43.3. The cut-off score for passing was 44. The reliability coefficient explains that 58% of the variance of test scores is nonerror, or 'true score" variance, while 42% of the variance is due to error; for example, in item selection or test situation (Cronbach, 1970, p. 165). In District C, where approximately one-half the students were above and one-half below the cut-off score, the reliability of .58 would result in making wrong decisions in approximately 22 percent of the cases (see Ebel, 1965, p. 334). That is, 11% of those failed actually passed according to their true score. By comparison, in District D, the same decision would result in only 7% error (the reading subtest had an estimated reliability of .93).

The variability in the reliability of the proficiency test in the seven districts ranged from .58 (District C, Reading) to .95 (District G, Mathematics), with a median reliability coefficient of .86 across the subtests. Given the importance of the decisions made on the basis of MCT tests, this low degree or precision cannot be tolerated. In our example of District C, 11% of eighth graders will be programmed in ninth grade into remedial classes for reading erroneously, which reduces the number of genuine English courses to

which these students will be exposed. Conversely, another 11% will be told they passed the test, when a more reliable test should have resulted in failure. These students will be denied remediation when they may truly need it. High schools in this district will have made "mistakes" in programming in ninth grade for more than one in every five students.

When we consider that most tests are designed to provide information about achievement of students in the vast middle range, the content is often not valid for children outside this middle range. Nunnally (1967) has noted that problems of test reliability are also related to range; tests developed for the middle level of learners at specific age and grade levels are not reliable for subjects outside these age and grade levels, since test scores become unreliable at the extreme ends of the score distribution.

Reliability Issues with Handicapped and Marginal Students

Thus far, the discussion is based on MCTs for a nonselect group of students, but the issues are further confounded for mildly handicapped and educationally marginal students. Provisions are made in most states for handicapped students, including the possibility of taking the MCT under modified procedures. What do the modifications do to the validity and reliability of the tests? It seems incongruous that a student can be declared competent or proficient in "reading" after passing a test read to him or her by the examiner. Similarly, the fact that modifications are permitted only for mildly handicapped students suggests that the group is rather homogeneous, a factor known to reduce the reliability of the test (Mehrens & Lehmann, 1984).

Comparability of Standards

Although the issue of standards for passing MCTs has been debated widely

(e.g., Haney & Madaus, 1978), the issue is compounded in local option states. The reason is that, in local option states, the matter involves not only what "level" of reading performance should be mandated, but rather whether the levels of proficiency in reading, for example, between districts are comparable in terms of grade equivalence required for passing. We illustrate the issue with data from the same seven school districts for which reliability data were presented.

Illustrative Data on Standards for Seven Districts

To compare standards for passing MCTs in the districts, a regression analysis was performed on the raw scores on the subtests of the MCT (e.g., reading) and the subtest raw scores of the standardized achievement test used in that district measuring similar content. For example, the reading scores on a district MCT were regressed on total reading scores from the Comprehensive Test of Basic Skills, the MCT math score on the Stanford Achievement Test total math score, and so on. Table 1.2 contains the grade equivalent score on the standardized achievement test corresponding with the cut-off score for passing the MCT on the corresponding subtest. Examination of Table 1.2 reveals that "minimum competence" is operationalized to mean achievement that ranges from that level expected of a beginning fourth grader (District G, Writing) to that expected of a senior in high school (District A, Math) and virtually every value in between. The variation in grade equivalents corresponding to cut-off scores is very great both across districts and within subtest areas and within districts and across subtest areas.

Table 1.2

Grade Equivalent Score on Standardized Achievement Test Corresponding to Cutting Score on District MCT on Corresponding Subtests

District	Reading	Writing	Mathematics
A	8.70	8.70	12.10
B	5.10	6.60	7.60
C	8.60	9.55	8.00
D	7.90	6.30	11.50
E	7.00	6.90	8.80*
F	8.30	7.40	8.60
G	5.50	4.00	10.00

Note. Ranges for Reading, Writing and Mathematics are 8.7 - 5.1 = 3.6; 9.55 - 4.0 = 5.5; 12.1 - 7.6 = 4.5, respectively. Median levels for Reading, Writing and Mathematics are 7th gr., 9 mos.; 6th gr., 9 mos.; and 8th gr., 8 mos., respectively.

Comparability Across Districts

The range of scores across districts in reading suggested that "minimum competence" ranges from reading at the grade equivalent score level of fifth grade, first month to eighth grade, seventh month (3 years and 6 months difference), while the ranges for writing and mathematics revealed differences of five years, five months and four years, five months, respectively, between the lowest and highest GEs corresponding to passing scores. Clearly, the standard for passing the MCT in any of the three areas mandated is not

constant; rather, it varied substantially among these seven districts. It should be noted that this variability is not an artifact of the standardized achievement test used, since the lowest and highest scores required for passing were found in districts using the same level of the same achievement test series.

Comparability Across Content Areas

The variability of standards for passing should also be noted across content areas within the same district. In District G, students are declared competent and permitted to graduate from high school when they demonstrate writing skills at about the fourth grade level, yet they must demonstrate math skills typically mastered in tenth grade. Every school district adopted standards (in terms of grade equivalence) that differed by "at least" 1 year for passing subtests of the MCT, and district G adopted standards that differed by 6 years.

Table 1.2 also reveals clearly that, in general, mathematics subtests are more difficult than are subtests in reading or writing. In fact, the standards in Districts A, D, and G are so high that they preclude high school graduation of many, if not most, educationally marginal students and mildly handicapped students unless drastic modifications in testing procedures are used or the teachers directly teach the test. It is instructive to note that Semmel, Gottlieb, and Robinson (1979), reviewing the research on mainstreaming, were unable to find a mean grade equivalent in reading achievement for mildly retarded students (including some with IQs to 85 and above) as high as 4.0 (p. 237). Granted, many of the children were not in 12th grade; yet the probability of current mildly handicapped students or those who are educationally marginal achieving at a level required to pass MCT subtests is

slight.

Summary

In the previous sections, we have illustrated our concerns about reliability of, and standards for passing, MCTs. High or low standards for passing MCT subtests are neither inherently good nor bad. Clearly, a high standard will result in more failures, greater remedial course enrollments, and more limited exposure to standard content courses and electives. Conversely, a lower standard leads to fewer remedial offerings and an increase in enrollments in standard content and elective courses. One must, however, consider standards in light of the reliabilities of the subtests. For example, District A had the highest standards for passing in reading and mathematics and was second highest in writing (see Table 1.2); yet all three of its MCT subtests have the lowest reliabilities of the seven districts. This district is making errors in determining who passes in nearly 20% of the cases!

It is our belief that local option competency testing has resulted in MCTs that too often lack the psychometric properties one would expect, given their important role in determining the educational careers of students. Given that the reliabilities of the subtests are undoubtedly lower than those reported for students at the lower end of the achievement distribution, we question the utility of these tests for making programming decisions. Furthermore, the standards for passing MCT subtests are so variable that we believe that districts frequently do not know the standards they are setting for their students. Even with modified testing procedures, few LD or EMR students will likely meet the standards in most of the school districts studied.

Conceptual Issues with Mildly Handicapped and Educationally At-Risk Students

One manifestation of the push for excellence in American educational has been the mandating of the use of minimum competency tests (MCTs) on a state-wide basis in a majority of states (Pipho, 1980). The tests and the uses to which they have been put (e.g., early warning of problems, requirements for graduation) have been debated extensively in the professional literature and the public forum. How MCTs should be used, if at all, with handicapped students continues to pose thorny problems (Cohen, Safran, & Polloway, 1980; McCarthy, 1983b; Smith & Jenkins, 1980).

Several authors (Cohen et al., 1980; Hall, Griffin, Cronin, & Thompson, 1985; Smith & Jenkins, 1980) have noted that the extension of MCT requirements to handicapped students was done without thoughtful consideration of the consequences of that policy. Will these tests prove beneficial or detrimental to the quality of education provided to handicapped students? It is our position that the potential for mischief currently exists, and that a number of conceptual issues must be addressed if MCTs are to be used as criterion for high school graduation. Furthermore, the discussion should include consideration of not only mildly handicapped learners (e.g., LD, EMR) but also that segment of the school population previously, but no longer, eligible for special education services as mildly handicapped learners.

Issue 1: Reciprocity and Competence as a Function of District

One issue that arises because of the differential passing standards for MCTs among districts is that of reciprocity and certifying the competence of at-risk students. This issue can be illustrated with hypothetical data based on the grade equivalent (GE) for a standardized achievement test corresponding

to the cutoff score for passing the MCT on the corresponding subtests. This can be accomplished through regression analysis performed on raw scores for the subtest of the district's MCT (e.g., reading) and the subtest raw scores of a standardized achievement test measuring similar content. (See Balow et al., 1986, for details of how this is done.)

Consider the following situations for two educationally at-risk 12th grade students who are functioning at a 7th grade level in writing. The student attending a high school in District #1 will have passed the MCT writing subtest and be granted a diploma, while the students in District #2 will be unable to pass the writing subtest in that district and will be denied a diploma (see Table 1.3). Given the variability in standards for passing, "competence" or "proficiency" in each subject matter area is a function of the district/test. McCarthy (1983b) alluded to this problem when she wrote: Thus, proficiency standards for high school graduation might differ throughout a state, posing a problem for students who transfer between districts. A student considered competent in a given subject areas in one district may not be able to fulfill the requirements in another district" (p. 108).

Table 1.3

Profile of Grade Level of Two Students

Student	District	Reading	Writing	Mathematics
#1	A	8.70	8.70	12.10
#2	B	5.10	6.60	7.60

McCarthy's concern is not simply a possibility; it is a reality. Most districts have a policy of reciprocity that recognizes passage of subtests in

other school districts. Using the above example, a student enrolled in District #2 passes the reading subtest of the MCT, achieving a GE of 6.0. This student then moves in District #1 and, given the reciprocity policy, is credited with passing the reading subtest of the MCT. Yet students who have been enrolled continuously in District #1 achieving at an even higher level (GE between 6.0 and 8.9) presumably fail the reading subtest. Hence, in this situation students who are "more proficient" in reading are judged "incompetent," while a student transferring into the district is judged "competent" by virtue of having passed a test that had a lower standard for passing in the previous district.

Consider a similar situation, when two students have identical scores on all three subtests. To illustrate, take students scoring at a GE of 7.0 on all three subtests -- reading, writing, and math. Using the standards for passing the various subtests shown in Table 1, consider the following situation. Student A is initially enrolled in District #2 and takes the MCT, passing the subtests in reading and math, and then moves to District #1. Student B is initially enrolled in District #1, but he moves to District #2 before the MCTs are administered in District #1. Student A is credited with having passed the subtests in reading and math, and on subsequent testing passes the writing subtest in District #1. This student, with 7.0 scores in all three areas, gets a high school diploma. Student B never gets a diploma because of the failure to pass the writing subtest in District #2.

Given the reciprocity policy and the different standards for defining proficiency, parents would be advised to "shop" for a district having standards for passing which their child can meet. This has the unfortunate consequence of directing the parents' (and the child's) attention toward how

to finesse the situation and away from the real problem -- remediation of the student's academic problems.

Such paradoxes are not limited to local option testing states, as reciprocity between states with different standards for passing leads to the same dilemma. As noted by Glass (1979), establishment of standards for passing MCTs is an arbitrary venture. Beyond this arbitrariness, however, the issue of rationality introduced by policies of reciprocity needs further consideration.

Issue 2: Rationality of Accommodations in MCT Testing for Handicapped Students

McCarthy (1983b) stated that excusing handicapped students from MCT requirements and granting them diplomas based on meeting IEP objectives ". . . might compromise the legal rights of the non-handicapped" (p. 108). This concern can be extended to permitting the use of modified testing procedures with handicapped students. Palardy (1984) wrote: "In many states, for example, students classified as educable mentally retarded (EMR) do not have to meet minimum competency requirements. But students whose capabilities are judged to be slightly greater than the maximum allowed for EMR classification are not exempt. How significant the difference between a 69 and a 71 IQ" (p. 404).

Contrast this passage with one from Smith and Jenkins (1980): " Many mildly handicapped youngsters (e.g., educable mentally retarded, learning disabled, emotionally disturbed) have, almost by definition, had difficulty taking standardized tests" (p. 440).

Some have argued that administration of MCTs under standardized test procedures to mildly handicapped students ignores the specific learning

problems that led to certification. That is, failure of MCT would confirm the diagnosis. Less compassion, however, is evident for the educationally at-risk students, many of whom would have been eligible for special education under more inclusive definitions of mental handicap that were employed during the 1960s (see MacMillan et al., 1988; MacMillan, 1989).

Permitting mildly handicapped students to either take MCTs under modified conditions or be judged competent under differential standards suggests that the "developmental vs. difference controversy" (Zigler & Balla, 1982) has been resolved in favor of the difference position. Do mildly handicapped children differ from educationally at-risk children in "degree" or "kind"? The position taken vis-a-vis MCT testing accommodations suggests the latter, yet the rationality of these policies can be questioned.

Let us consider first the use of differential standards, which permits special education students to meet the MCT requirement by completing IEP established objectives. Should educationally at-risk students, by this same reasoning, be considered "competent" by virtue of successful completion (e.g., passing grades) of courses required for graduation despite being unable to pass all substest of the MCT? The San Diego Unified School District has dropped its single test for assessing competence and has assigned the required competencies to passage of those courses required for graduation. By virtue of securing passing grades in each required course, the student is deemed to have passed the minimum competency requirement (Singer & Balow, 1987). At present, this alternative is not open to educationally at-risk students, yet the course of study prescribed for nonhandicapped students seems analogous to an IEP for a handicapped student.

The most paradoxical situation arises where an educationally at-risk

student is denied a diploma because of failure on MCT subtests and yet is "more competent" (i.e., on a comparative scale such as measured achievement) than a handicapped student who meets the MCT requirement through differential standards. One major goal of the MCT movement is to have a high school diploma "reflect achievement rather than social promotion" (McCarthy, 1983b, p. 105), this goal is abrogated by the provision of differential standards of testing. Diplomas will be granted, in some instances, to special education students whose achievement is substantially lower than that of students being denied diplomas by the same school district. As this may be construed as a form of social promotion, the rationality of this state of affairs must be questioned.

The use of alternative modes, or modified testing procedures, appears to be a logical, yet unstudied, means of accommodating the specific learning problems of handicapped students. Cohen et al. (1980) noted that a child who has difficulty reading written materials but who is able to comprehend orally presented material might be accommodated with an auditory presentation. This is particularly true for a child for whom this same accommodation has been made instructionally; that is, the MCT should not consist of items presented in a manner different from that by which the student has been instructed or tested in the past. The purpose of modifying the testing procedure for handicapped students is to measure the proficiency in the subject-matter area free of any interference from poor test-taking skills.

What has not been studied is the effect of the various modifications on the reliability and difficulty of the tests. That is, do certain types of modifications make the results less reliable or difficult? Although it may seem logical to present the reading subtest orally to a student who has a

specific problem decoding printed material, this form of modification could affect the reliability of the test and might make the test easier for any student, regardless of handicap. If that were the case, then all students should have the option of using a particular alternate mode, because the alternate mode should only enhance their level of performance. Such a choice would be most salient for educationally at-risk students, given their greater likelihood of failing MCT subtests administered under standard conditions. This study represents an attempt to evaluate the effects of the alternative modes being used on reliability and difficulty of the MCTs.

Issue 3: Curricular Consequences of Failing MCTs for At-Risk: Equality of Educational Opportunity?

MacMillan and colleagues (1988) traced the pressures that mounted during the 1960s and 1970s to modify special educational designations under the belief that a significant proportion of mildly handicapped (e.g., EMR, LD, BD) children would be better served as unlabeled students in regular classes. Through litigation and other means, the definitions of EMR and LD have been made more exclusive in the belief that many children heretofore labeled EMR were denied equal educational opportunity. That segment of students, however, no longer eligible for special education (under a press for equality of opportunity) is now going to be held to MCT standards (i.e., press for excellence) they are ill-equipped to meet (see Balow et al., 1986).

The educationally at-risk segment of the school population will, in considerable numbers, fail the MCT subtests when taken initially (usually 8th or 9th grade). For those students who fail subtests, districts must provide remediation. Frequently this remediation takes the form of enrolling students in remedial math (or writing or reading) course during the subsequent school

year. The result of this remedial model will be that educationally at-risk students will take mandated courses for graduation, with "elective spots" in their program filled with the remedial courses resulting from MCT failure. If the MCT subtests are repeatedly failed (e.g., in 9th, 10th, and 11th grade), a given student could conceivably complete four years of high school without an elective. This would probably lead to perceptions of the high school curriculum as irrelevant, since elective courses of interest cannot be taken because remedial courses have filled these slots.

The long-range consequences of this development for the high school curriculum are considerable. The demand for nonacademic/nonremedial courses (e.g., vocational ocurses) will diminish. Students for whom higher education is unlikely (the educationally at-risk group) will be denied the opportunity to take the very courses that would increase their employability and possibly keep them interested in school. Rather, they will take a program of preparation almost exclusively devoted to college preparatory courses plus remedial courses in academic tool areas. To what extent are MCTs implicated in the escalating dropout rate in high schools? The Committee on Policy Research (1985) of the California Assembly reported the following: "Students drop out because they are not succeeding in school, they want to work, or they are pregnant . . . Some dropouts fail district proficiency tests or courses required to graduate. Of the 100,000 dropouts in the class of 1983, 41,000 left in 12th grade due to failure of proficiency tests or the courses needed to graduate" (p. iii).

Failure of MCTs by handicapped students can be accommodated through modified testing or use of differential standards. The majority of regular class students will pass all subtests by the early high school years. The

educationally at-risk segment of the student body, however, is composed of likely candidates for repeated failure of MCT subtests, resulting in severe restrictions on elective courses to be taken. To the extent that the old EMR category overlaps with current educationally at-risk populations, it seems likely that lesser, rather than greater, educational opportunity was achieved as a result of the court challenges of the 1970s. The descriptor "educational dead-end" used by Judge Peckham to describe EMR classes of the 1970s in his opinion in Larry P. (1971) more aptly describes the curriculum to which the current educationally at-risk student is forcibly exposed, rather than the 1970s EMR program.

Remedial Classes: Revisitation of the Debate Over Special EMR Classes

Thus far, the data reported by states suggest considerably higher failure rates for minority students than for majority students. California reported for the year 1982 that 9992 black students, 13,567 Hispanic students, and 47,262 white students completed proficiency exams during the 12th grade. The percentages failing one or more subtests were as follows: 35% of the black students, 29% of the Hispanic students, and 15% of the white students. Serow (1984) reported on students initial efforts with MCTs in California, Florida, North Carolina, and Virginia. In California, the combined reading and math tests resulted in passage rates for whites and blacks, respectively. The greatest discrepancy appeared on math subtests in Florida and North Carolina, where 83% and 95% of whites passed, while only 40% and 73% of blacks passed. Serow (1984) reported that blacks constituted 27% of the "Class of 1980" in North Carolina and yet accounted for 76.1% of all diploma denials due to MCT failure. The differential impact of MCTs on blacks and whites has led many (e.g., Cawelti, 1978; Fisher, 1978; Pullin, 1982) to question the fairness of

the MCTs.

The parallels between the concerns expressed over EMR classes in the 1960s and 1970s and those arising over MCT testing are considerable. First, racial differences in "passing" rates on tests could raise questions of test bias. The intelligence tests challenged during the 1960s and 1970s were norm-referenced, while current MCTs are criterion-references tests. The implicit position taken by critics of both types of tests is that, in order for a test to be fair, it must yield similar means and variances for different groups. Existing MCTs clearly do not yield similar means and variances for students from different racial or ethnic groups, so concerns regarding test bias in MCTs (criterion-referenced tests) must surely be as reasonable as similar concerns regarding intelligence tests (norm-referenced tests).

Second, the EMR classes were criticized during the 1960s and 1970s for containing a disproportionately high percentage of minority students. MCT testing is resulting in a disproportionately high failure rate for minority group students, which will inevitably result in remedial classes with disproportionately high minority enrollments. If the goal of the court actions regarding EMR classes was to eliminate the "de facto" segregation of minority children, the goal has proven most elusive -- present-day remedial classes pose the very same problem the court rulings were designed to eliminate.

Third, one impetus for changing the EMR special class was concern over stigma associated with labeling and segregation. As Dunn (1968) wrote regarding disability labels, "we must examine the effects of these disability labels on the pupils themselves. Certainly none of these labels are badges of distinction. Separating a child from other children in his neighborhood -- or

removing him from the regular classroom for therapy or special class placement -- probably has a serious debilitating effect upon his self image (p. 9)".

Excepting the label, the process of assigning those students who fail MCTs to remedial classes seems virtually identical to the process described above. Ironically, the group of children over which the litigation raged in the 1970s in the name of equality of educational opportunity (minority children achieving academically at a marginal level) is the very group exposed to the very same process in the 1980s in the name of promoting excellence in American schools.

Summary

This examination of conceptual issues surrounding the use of MCTs with mildly handicapped and at-risk students underscores what little consideration was given to the extending MCT requirements to handicapped students (Cohen et al., 1980; Hall et al., 1985; Smith & Jenkins, 1980). Permitting modifications in testing procedures for handicapped children has not been systematically evaluated and may alter the reliability and difficulty of the tests. If so, the rationality of permitting such modifications for mildly handicapped children while denying their use with at-risk students must be questioned. These two groups differ in degree, not kind, and using arbitrary cut-offs for IQ and achievement to determine which students get modified testing and which do not seems difficult to justify. Granting diplomas to handicapped students on the basis of differential standards and denying diplomas to more proficient at-risk students who are unable to pass MCTs under standardized testing conditions is similarly mischievous.

The concern so apparent in the 1970s concerning the educational plight of students in the IQ range of 70 to 85 is not apparent today as those in this

IQ range struggle to obtain a diploma. The change in definition of mental retardation was prompted by the belief that the upper echelon of the "old EMR" group were better served in regular programs. Subsequent developments such as the MCT movement call into question whether this difficult-to-teach segment of our student population is better served in regular programs. It is apparently that competency or proficiency is elusive, as one considers the variability in standards for passing (Balow et al., 1986); there is no evidence that MCTs have improved the quality of instruction or the level of performance for mildly handicapped learners. Moreover, MCTs appear to serve as a hindrance to the students termed educationally at-risk, and the irrationality in the way they are implemented leads to the call for a prompt and serious re-examination of this requirement.

References

Anderson, B. D. & Lesser, P. (1978). The cost of legislated minimum competency requirements. Phi Delta Kappan, 59, 602-608.

Balow, H. B., MacMillan, D. L., & Hendrick, I. G. (1986). Local option competency testing: Psychometric issues with mildly handicapped and educationally marginal students. Learning Disabilities Research, 2(1), 32-37.

Brickell, H. M. (1978) Let's talk about minimum competency testing. Denver: The Education Commision of the United States.

Candor-Chandler, C. (1978). Charleston, West Virginia: Competency requirements for special education students. Phi Delta Kappan, 59, 610-612.

Cohen, S. B., Safran, J., & Polloway, E. (1980). Minimum competency testing: Implications for mildly retarded students. Education and Training of the Mentally Retarded, 15, 250-255.

Committee on Policy Research (1985, August 23). The Dropout Problem in California. Sacramento, California.

Cronbach, L. J. (1970). Essentials of psychological testing. New York: Harper & Row.

Diana v. State Board of Education (1970). Civil Action No. C-70-37 (N.D. Cal., 1970).

Dunn, L. M. (1968). Special education for the mildly retarded: Is much of it justifiable? Exceptional Children, 35, 5-22.

Ebel, R. L. (1965). Measuring educational achievement. Englewood Cliffs, NJ: Prentice-Hall.

Fisher, A. T. (1978). Four approaches to classification of mental

retardation. Revised version of paper presented at A.P.A. Annual Convention, Toronto, Canada.

Forness, S. R. (1985). Effects of public policy at the state level: California's impact on MR, LD, and ED categories. Remedial and Special Education, 6(3), 36-43.

Glass, G. V. (1978). Minimum competence and incompetence in Florida. Phi Delta Kappan, 59, 602-605.

Glass, G. V. (1979). Looking at minimal competency testing. Education and Urban Society, 12, 47-55.

Gorth, W. & Perkins, M. (1979). A study of minimum competency programs: Final summary and analysis report. Amherst, MA: National Distribution Systems.

Grossman, H J., (Ed.) (1973). Manual on terminology and classification in mental retardation. Washington, D.C.: American Association on Mental Deficiency.

Hall, J., Griffin, H., Cronin, M., & Thompson, B. (1985). Factors related to competency test performance for high school learning disabled students. Educational Evaluation and Policy Analysis, 7, 151-160.

Haney, W., & Madaus, G. (1978). Making sense of the competency testing movement. Harvard Educational Review, 48, 462-484.

Hart, G. K. (1978). The California pupil proficiency law as reviewed by its author. Phi Delta Kappan, 59, 592-595.

Jaeger, R. M. (1982). An interative structured judgement process for establishing standards on competency tests: Theory and application. Educational Evaluation and Policy Analysis, 4, 461-475.

Kelly, T. (1927). Interpretation of Educational measurements. Yonkers, NY:

World Book.

Klein, K. (1984). Minimum competency testing: Shaping and reflecting curricula. Phi Delta Kappan, 65, 565-567.

Lambert, N. M. (1981). Psychological evidence in Larry P. v. Wilson Riles: An evaluation by a witness for the defense. Exceptional Children, 54(4), 297-301.

Larry P. v. Wilson Riles (1971). Civil Action No. 71-2270 (N.D. Cal., 1971).

Lewis, D. (1979). Certifying functional literacy: Competency testing and implications for due process and equal educational opportunity. Journal of Law and Education, 8, 145-148.

Logar, A. (1984). Minimum competency testing in schools: Legislative action and judicial review. Journal of Law and Education, 13(1), 35-49.

MacMillan, D. L. (1988). Issues in mild mental retardation. Education and Training in Mental Retardation, 23, 273-283.

MacMillan, D. L. (1989). Equality, excellence, and the EMR populations: 1970-1989. Psychology in Mental Retardation and Developmental Disabilities, 15(2), 1, 3-10.

McCarthy, M. M. (1980). Minimum competency testing and handicapped students. Exceptional Children, 47, 166-173.

McCarthy, M. M. (1983). Minimum competency testing for students: Educational and legal issues. Educational Horizons, 61, 103-110.

McCarthy, M. M. (1983b). Minimum competency testing for students: Educational and legal issues. Educational Horizons, 61, 103-110.

McKinney, J. D. (1983). Performance of handicapped children on the North Carolina minimum competency test. Exceptional Children, 49, 547-549.

Mehrens, W. A., & Lehmann, I. J. (1984). Measurement and Evaluation in

Education and Psychology. New York: Holt, Rinehart, & Winston.

Nunnally, J. C. (1976). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Olsen, K. R. (1980). Minimum competency testing and the IEP process. Exceptional Children, 47, 547-549.

Palardy, J. M. (1984). An analogy for reviewing promotion/retention ns minimum competency decisions. Education, 104, 410-404.

Pipho, C. (1978). Update VIII: Minimal competency testing. Denver, CO: Educational Commission of the States. May.

Pipho, C. (1980). State minimum competency testing programs: Analysis of state minimum competency testing programs. Final report. Denver: Education Commission of the States. Department of Research and Information Services (65 pages).

Polloway, E. A., & Smith, J. D. (1988). Current status of the mild mental retardation construct: Identification, placement, and programs. In M. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), Handbook of special education: Research and practice. Vol II (pp. 7-22). Oxford, England: Pergamon Press.

Public Law 94-142, Education for all handicapped children act, November 29, 1975.

Pullin, D. (1982). Minimum competency testing, the denied diploma and the pursuit of educational opportunity and educational adequacy. Report. Grant NIE CSE-R-180. Center for the Study of Evaluation, University of California, Los Angeles.

Safer, N. D. (1980). Implications of minimum competency standards and testing for handicapped students. Exceptional Children, 46, 288-290.

for handicapped students. Exceptional Children, 46, 288-290.

Semmel, M.I., Gottlieb, J., & Robinson, N. M. (1979). Mainstreaming: Perspectives on educating handicapped children in the public schools. In D. Berliner (Ed.), Review of Research in Education. (Vol. 7) (pp. 223-279). Washington, D.C.: American Educational Research Association.

Serow, R. C. (1984). Effects of minimum competency testing for minority students: A review of expectations and outcomes. The Urban Review, 16(2), 67-75.

Singer, H., & Balow, I. H. (1987). Proficiency Assessment and Its Consequences. California Policy Seminar, Institute of Governmental Studies, University of California, Berkeley.

Smith, J. D., & Jenkins, D. S. (1980). Minimum competency testing and handicapped students. Exceptional Children, 46, 440-443.

Walstad, W. B. (1984). Analyzing minimal competency test performance. Journal of Educational Research, 77, 261-266.

White, A. O. (1984). Florida's functional literacy test on trial. Urban Education, 19 (1), 17-28.

Wise, A. E. (1978). Minimum competency testing: Another case of hyperrationalization. Phi Delta Kappan, 59, 596-599.

Zigler, E., & Balla, D. (Eds.) (1982). Mental Retardation: The developmental-difference controversy. Hillsdale, NJ: Erlbaum.

CHAPTER II

Design and Methods

The investigation reported herein compared three categories of subjects: (a) Learning Handicapped (LH) students, (b) Educationally Marginal (EM) students, and (c) Regular Controls (RC). Initially, we describe each of these groups and explain the rationale for including these three groups in the study of competency testing. We then explain how samples of each of the three groups were selected, and finally describe the data collected on each subject.

Subjects

In 1974, the California Master Plan (California State Department of Education, 1974) abolished the traditional categorical grouping of handicapped children. Prior to that time, California's special education programs differentiated programs for children categorized as educable mentally retarded (EMR) and Educationally Handicapped (EH), the latter being a broader rubric than those used in most other states. Included under the EH rubric were behaviorally disordered (BD), children with suspected minimal brain dysfunction (MBD), emotionally disturbed (ED) children, and those who in other states would be classified as learning disabled (LD). Hence, mildly handicapped students in the California schools prior to the California Master Plan were categorized as either EMR or EH. The California Master Plan abolished the EMR and EH categories and created a new, and broader, category of mildly handicapped learners called Learning Handicapped (LH) which subsumes children heretofore classified as EMR or EH in California. In essence, the LH category serves children/adolescents with mild learning handicaps and does so in either special day classes (SDC) or resource specialist programs (RSP). Students served in LH programs are system-identified special education

students certified by school districts as eligible for, and enrolled in, special education programs. According to the California Education Code, Learning Handicapped students are defined to include:

> ...pupils under the age of 21 years, who by reason of marked learning or behavior disorders, or both, cannot profit from the regular education program, and who, as a result thereof, require special education programs. Such learning or behavior disorders can be associated with a neurological handicap, emotional disturbance, learning disability, mild handicap, or can be attributable to mental retardation.

Our interest in the LH population derives from the fact that since they are designated as in need o special education services, these students are provided with avenues for passing the MCT requirement in addition to the standard passing of the test, as described in Chapter 1. The possibility of graduating from high school does not, for special education students, hinge entirely on the ability to pass the MCT under standardized testing conditions.

The second group of interest are herein called "Educationally Marginal" (EM) students. These students are enrolled in regular classes and are regular students in the sense that they are not system-identified as eligible for special education services. The students we define as EM are those whose level of academic achievement places them at the extreme lower end of the achievement continuum. We anticipated, and indeed found, among the larger EM group a substantial number of students whose achievement level, and measured general intelligence, suggest they resemble the "borderline mentally retarded" classification included in the Heber (1961) definition whose IQs were in the range of approximately 70-85. Following court cases (e.g., Larry P. v.

Riles), and legislation (e.g., P.L. 94-142), and the adoption of a more exclusive definition of mental retardation (Grossman, 1973), the cohort of children with IQs between 70 and 85 were no longer eligible to be classified as EMR. Moreover, these students (IQ 70-80 or 85) cannot qualify for services as LH via the "LD route" according to a learning disability "discrepancy formula" published in a manual by the California State Department of Education (1983). Forness (1985) referred to this group of students (i.e., those in the IQ range of 70 - 82) as falling in special education "no man's land" as the published tables for establishing cut-offs essentially preclude designating a child with an IQ lower than 85 as "learning disabled" regardless of how low the child's achievement might be. The tables for establishing cut-off scores published by the State Department of Education, utilizing the WISC-R and either the PIAT or the WRAT, show the lowest IQ bands which could be considered for a discrepancy determination to average 82.7. Hence, children with IQs between 70-85 are ineligible for the LD-route to LH because their low achievement (when it occurs) is attributed to low aptitude. These same students are ineligible for the "EMR-route" to LH because their IQs are too high to certify them as mentally retarded. Hence, they are normal students by default.

We proposed to study this EM group due to the possibility that they contain a number of "unserved handicapped" students according to P.L. 94-142, which includes the following:

> ...there are many handicapped school children throughout the United States participating in regular school programs whose handicaps prevent them from having successful education experience because their handicaps are undetected. (Section 601)

This concern is possibly realized in California under eligibility guidelines operating when this project began, as Forness noted in the following:

> There no doubt continues to be a group of children possibly in the 70 or 75 to 80 IQ range and very probably with accompanying poor adaptation to the social and academic demands of school, who are not in special classrooms nor are they receiving resource assistance in regular classrooms. (Forness, 1985, p. 38)

This EM group is required to pass MCTs in order to receive a high school diploma. Furthermore, the EM students must pass the MCT under standardized testing conditions, since they are not certified eligible for special education. That is, the modifications in MCT testing available to special education students cannot be made available to EM students because they are nonhandicapped, or normal, students.

The final group of students, Regular Controls (RC), constitute those whose measured achievement scores place them above the 25th percentile. As such, they provided us with baseline data against which the performance of the LH and EM groups' performance could be interpreted.

Sample Selection

The sampling plan employed stratified random sampling in the selection of LH, EM, and RC subjects. Within each of the sample categories, subjects were stratified on the basis of gender and ethnicity (Anglo, Black, and Chicano) and then selected at random until sufficient cases were secured. Four school districts in southern California cooperated in the project for the entire five-year period of the project. While our initial plan called for selecting equal numbers of subjects in each cell (e.g., Anglo-Male-LH), the

actual available cases within the cooperating schools made this impossible as available cases for a given cell were too few. In fact, there were so few Black children in LH programs we modified the design and used a cohort design with one Cohort being selected from 8th graders in 1985, and a second Cohort being selected from 8th graders in 1986.

Learning Handicapped (LH). We secured from each of the four cooperating school districts a list of all 7th graders certified as eligible, and being served, as LH the year prior to the commencement of the study. These lists were then separated according to the gender and ethnicity of the students (e.g., a separate list for Black-Female-LH, etc.) and randomly ordered. Informed consent forms were then sent to parents/guardians of selected children. These forms were to be returned to the school. If the child failed to return the consent form within 10 days, a second form was mailed to the child's home and follow-up phone calls were made by project staff.

The procedure used in the selection of Cohort II was almost identical to that used in the selection of Cohort I LH subjects. One bilingual (Spanish and English) project employee phoned all homes when Spanish was indicated as the primary language, or when other project personnel were unable to communicate with family members.

In the selection of Cohort II subjects, the procedures were somewhat modified in two of the four districts. In these two districts, eighth grade students were given information and consent forms at the school site. Students were instructed to take the forms home and return the bottom portion (the consent) indicating parental consent. All returned forms were included in an incentive drawing for prizes. These prizes were donated by local fast food restaurants and included items such as free drinks, pizzas, lunches,

pencil/eraser sets, etc. Students understood that a signed consent form did not guarantee participation in the project, as random selection was to be used to select those from the total population.

Educationally Marginal (EM). Scores from standardized achievement tests of all students who were in the 7th grade the year prior to the study were secured from each of the four districts. Scores were available for Reading/Language and Mathematics subjects from the previous Spring administration of standardized achievement tests. Those students whose test scores placed them in the lowest quartile in either Reading or Math were placed on a separate list, as potential EM subjects. These lists were then stratified by gender and ethnicity. The final stratified lists were then randomly ordered and consent letters sent to the homes in the same fashion as was done for LH students. Follow-up letters and phone calls were also used.

Once informed consent was received a final list of names for EM subjects was compiled in an effort to approximate the ethnic and gender composition of the LH sample. Once potential EM students were identified and consent received, project personnel administered the Otis-Lennon School Ability Test (Otis & Lennon, 1982) in small groups. Following that, all potential EM students were administered individually the Slosson Intelligence Test (SIT) (Slosson, 1985). Performance on the SIT would serve as the basis for dividing the EM sample into "high" and "low" IQ subsets, with the low EM group approximating the former borderline mentally retarded category with IQs under 85.

Regular Controls (RC). After lists of all EM-eligible subjects (i.e., those scoring in the lowest quartile in either Reading or Math) were created, the remaining students on the 8th grade lists were stratified according to

gender and ethnicity, with the order randomized. A sample proportionate to the LH and EM sample, in terms of gender and ethnicity, was then selected. Informed consent letters were sent to the homes, with follow-up forms and telephone contact used when forms were not returned within 10 days. When consent was not given, replacements were selected from the appropriate lists, which had previously been randomly ordered, and consent forms sent.

The initial sample, of 8th grade students, consisted of three student achievement categories (LH, EM, RC) X three ethnic groups (Anglo, Black, Chicano) X two gender (male, female) groupings (total N = 1188). Some planned comparisons involved subdividing the EM sample at IQ 85 (above and below) to consider high and low subgroups within the EM sample. Table 2.1 shows the breakdown of the initial 8th grade sample by category, ethnicity, and gender and Table 2.2 shows the demographics of the second cohort by category, ethnicity and gender.

Table 2.1

Demographic distribution of Cohort I MCT Sample by Ethnic, Gender and Category

	Cohort I					
	RC		EM		LH	
	Female	Male	Female	Male	Female	Male
Anglo	21	87	26	67	12	53
Black	10	10	5	12	5	7
Chicano	11	29	24	24	7	14

Table 2.2

Demographic distribution of Cohort II MCT Sample by Ethnic, Gender and Category

	Cohort II					
	RC		EM		LH	
	Female	Male	Female	Male	Female	Male
Anglo	81	81	46	58	35	59
Black	53	53	10	18	6	13
Chicano	64	54	41	46	14	32

Procedures

Once subjects were selected according to the plans described above, project staff went to district office files and school building files to collect demographic data on the subjects participating in the project. This activity of collecting demographic and school-related data was ongoing during the first year of the project and was then performed during summer months between academic years for remaining years of the project. A sample form on which research assistants recorded data from files is shown in Appendix A. In addition, during the Spring of every year subjects were administered the Self-Description Questionnaire II (SDQ II), a self-concept scale for middle adolescents (Marsh & Barnes, 1982), and the Survey of School Attitudes (SSA), a scale for assessing attitudes toward school (Hogan, 1975). Copies of these surveys are found in Appendices B and C, respectively.

The SDQ II was administered at the school sites, in groups of approximately 30 students. Each of the 140 SDQ II items was read aloud twice to the students, who were instructed to select one of six responses (False, Sometimes False, More False than True, More True than False, Sometimes True, and True). Administration of the SDQ II took approximately 35 minutes for

each group of students. The SSA was also administered at the school site, to groups of approximately the same size. The sixty items on the SSA were read aloud twice to the students, who then selected one of three response alternatives (Like, Not Sure/Don't Care, or Dislike). Administration of the SSA took approximately 15 minutes for each group of students.

Instruments and Assessements

SDQ II. Self-concept was assessed with the Self-Description Questionnaire II, which is based on the theoretical work of Shavelson, Hubner, and Stanton (1976). The SDQ II is one of three self-concept measures developed by Marsh -- the SDQ for preadolescents, the SDQ II for early adolescents, and the SDQ III for late adolescents. The SDQ II scale used in this project consists of 140 simple, descriptive statements to which the subject indicates agreement-disagreement on a six-point Likert scale, varying from 1 = False to 6 = True. The instrument yields measures of eleven components of adolescent self-concept: General, Math, Verbal, Academic, Honesty, Parents, Emotions, Physical (i.e., physical abilities), Appearance, Same Sex, and Opposite Sex. Factor analyses have confirmed these dimensions of self-concept as comprising independent factors. Marsh, Parker, and Barnes (1985) reported that factor analysis clearly identified each of the eleven scales; each scale was reliable (median r = .86). Marsh (1986) further demonstrated the empirical discriminability of these eleven factors of self-concept yielded by the SDQ II through multitrait-multimethod analyses.

Data collected on the current project were subjected to a series of factor analyses utilizing the eighth grade SDQ II data. Data collected on samples stratified by academic level (RC, EM, and LH), by ethnicity (Anglo, Black, and Chicano), and gender were factor analyzed. These factor analyses

were conducted within each of the above groups to determine whether groups had similar factor structure, in order to ascertain whether comparisons on scale scores could be meaningfully interpreted. Across all groups of subjects, and levels of stratification, the hypothesized eleven-factor structure of the SDQ was strongly supported by all methods of assessment. Each of the hypothesized factors was well identified. The results of the factor analyses can be found in the report by Little, Widaman, Farren, MacMillan, Hemsley, and MacMillan (1990a).

Internal consistency reliability analyses were also performed on the eighth grade SDQ II data to validate the psychometric properties of the SDQ scales in the samples selected. Across 1140 students, (48 did not have complete self-concept data) coefficient alpha internal consistency reliabilities for the self-concept scales were quite high, ranging from .80 to .90, with a median of .86. For the 534 RC subjects, coefficient alphas for the scales ranged from .81 to .91, with a median of .86; for the 364 EM subjects, coefficient alphas ranged from .80 to .89, with a median of .85; and, for the 242 LH subjects, coefficient alphas ranged from .75 to .88, with a median of .78. Data on the RC and EM students revealed high and comparable levels of internal consistency reliability, whereas data from LH students had slightly lower, but still quite acceptable levels of reliability.

The concordance of structure across all the groups analyzed allows legitimate and valid comparisons between the groups based on relative self-concept scores yielded by the SDQ II. In light of problems previously encountered when self-concept scales were used with mildly handicapped students, the findings from the factor analyses were extremely encouraging with regard to the psychometric properties of the SDQ II when used with

subjects at the lower extremes of the achievement distribution. Scales developed for the general population are often found to have substantially lower reliabilities than generally reported when used with individuals at extremes of a distribution (Nunnally, 1976). Yet, internal consistency estimates for the LH and EM samples indicated that Academic and Nonacademic self-concept subscales of the SDQ-II were highly reliable for all groups (Little, et al., 1990a).

SSA. Attitudes toward school were assessed with the Survey of School Attitudes (SSA), developed by Hogan (1975). The SSA consists of 60 items, divided into four scales (mathematics, social studies, science, and reading/language). Subjects are instructed to select one of three responses (Like, Not Sure/Don't Know, or Dislike) in responding to each of the 60 items. The SSA treats school-related attitudes as multidimensional in nature, and evidence exists supporting the notion that students respond differentially to scales assessing different dimensions of school related attitudes (Beck, 1977).

Hogan (1975) reported split-half reliability and alpha coefficients on the SSA, which ranged from .77 to .91, with a median of .85. Based on a large national sample, norms are available by item and scale for each grade. Conversations with the test author (T. P. Hogan, personal communication, April 16, 1985) suggested that scores on the four scales are decreasing by junior high school age and continue to decrease during the high school years. Because the normative scores on most items and scales at the 8th grade level are close to midrange, it was felt that there would be adequate range remaining to accommodate attitude changes for several years. Preliminary analyses were performed on data from the present study to investigate the

psychometric properties of SSA scales in our sample. Across all 1140 students, coefficient alpha internal consistency reliabilities for the four school attitude scales were quite high, ranging from .830 to .908. For the 534 RC students, coefficient alphas for the scales ranged from .817 to .912; for the 364 EM students, coefficient alphas ranged from .833 to .903; and, for the 242 LH students, coefficient alphas ranged from .838 to .910. Hence, data from all three groups of students revealed rather high and comparable levels of internal consistency reliability (Little, Widaman, Farren, MacMillan, Hemsley, & MacMillan, 1990b).

Additionally, factor analyses were conducted within each group to determine whether groups had similar factor structure for the SSA. This is an important question, as comparisons on scale scores are only meaningful if similar factor structures characterize the groups. In a series of factor analyses, data from samples stratified by academic level (three categories: RC, EM, and LH), by ethnicity (three groups: Anglo, Black, and Chicano), and by gender were factor analyzed. Across all forms of stratification, the four hypothesized factors were well identified. The results of the factor analyses, along with the reliability results described above are reported in Little et al. (1990b).

Standardized reading and mathematics achievement test scores. Test scores derived from standardized achievement tests in reading and math were secured from district-mandated achievement tests, administered each Spring. In this project, only scores in reading/English and math skills were utilized. Participating districts used different tests (e.g., 3Rs, Stanford Achievement Test, and the Comprehensive Test of Basic Skills); and different levels of these tests were administered to different students within the same district.

Districts also recorded scores in different formats (e.g., raw scores, percentile, etc.). Because of these many, and significant, variations in reporting achievement scores, it was necessary to convert scores to a common metric. This was accomplished by converting individual scores first to percentiles (as given in the test manual of each particular test and level) and subsequently to Normal Curve Equivalency scores (NCEs). NCEs yield a common metrict by which the measures of achievement are comparable between the four districts (Crocker & Algina, 1986). NCE scores have a mean of 50 and a standard deviation of 21. This standardization still maintains relative relations among individuals as well as differences in means and variances (as well as covariances) between the groups. At the stage of analyses, all standardized achievement test scores were in NCE form.

Minimum Competency Test data. MCT scores were gathered yearly from either cumulative records at the school site, or printouts generated by the districts' Data Processing departments. Scores from the math and reading portion of the MCT were averaged, forming the MCT score.

MCT scores were also converted to NCE format, for not only were the scores from four different school districts, but the versions (and sometimes, content) of the tests changed throughout the course of the study. As mentioned previously, the use of NCE scores assures a common metric.

Students frequently do not pass the test on the first attempt in the 7th grade, and some students take the test every year through the 12th grade before passing. As such, each year of MCT scores (in NCE format) were averaged to create one score used in the analyses.

An additional variable representing MCT status was created. This dichotomous variable indicated whether the student passed both the reading and

math portions of the MCT, or had failed at least one subtest.

Additional Variables

All 1188 students were accounted for at the time of analyses as being a transfer student, a drop-out, or a "regular" student (on-schedule for graduation). In addition to belonging to one of these categories, students were also classified into a dichotomous over-age variable.

Transfer & Drop-out. Designation as a drop-out or a transfer student (variable value of 1) indicates that a student is no longer in the original district or comprhensive high-school where he/she was originally enrolled. If students neither dropped out nor transferred, they articulated normally (variable value of 0). Because there are only three categories into which a student can fall, some district classifications (such as job corp, expulsion, or pregnancy) have been collapsed into one of the three groups. One subject died during the study so his ID number was removed from the sample. Table 2.3 summarizes how various student actions were classified as either drop-out or transfer behavior.

Table 2.3

Qualifications for Drop-out and Transfer Classifications

Drop-Out Classification	Transfer Classification
No-show	Out-of-District
Expulsion	Within District
Withdrawn	Private School
Juvenile Hall	Continuation School
Runaway	Independent Study
Over-age	Pregnancy Program
Job Corp	Graduate/GED
Unknown	Home Study

Over-Age. Over-Age was operationalized as those students who would turn 18 on, or before, November 15 of their senior year. The November cut-off date and the 18 year-old age was utilized because November is the date used by most districts in determining eligibility for Kindergarten. Turning eighteen before this November data represents one year beyond the age which students should complete high-school if they enrolled in Kindergarten at age five and were normally articulated throughout their academic career.

Although this variable does not indicate the reason for a students' older age (e.g., retention, illness, age of entry to kindergarten) it is a loose proxy of students' prior academic performance.

Course Selection. This scale variable was designed to determine the types of courses in which students enrolled. After consulting with school counselors and the high-school handbooks from each of the participating schools, courses taken by subjects were coded as being required, remedial, or elective. Courses were then coded with a 1, 2 or 3, respectively. These codes were refined to create a variable (CLS) representing an average of the types

of classes the student enrolled in during a given school year.

Discipline. A combination of the various types of discipline referrals (suspension, anti-teacher, student-conflict, maladjusted behavior, subtance abuse and unspecified) was used to create a comprehensive discipline variable for each grade. The resulting three-level scale variable indicates when a subject has received no referrals, 2 or less, and 3 or more referrals for any combination of the above behaviors.

Absence data. The average number of days a student missed school over the five years of the study were averaged to create the attend variable. This is a continuous variable ranging from 0 to 83.5 days absent.

Grades. Grades received in math and reading classes taken in 8th through 11th grades were averaged to create this GPA variable. Initial grades were converted as follows: 1 = F, 2 = D, 3 = C, 4 = B, 5 = A. It should be noted that this scale is a 1 through 5 scale, as opposed to the more common 1 through 4 scale typically used when calculating GPAs. Missing data were replaced using a general linear model procedure where the missing grade (dependent variable) was replaced by predicting its value based on grades in other classes and previous years' grades. This procedure was done by ethnicity, category, and gender.

Summary

Clearly, a variety of dimensions of a students' academic experience was considered in this investigation. The thoroughness of the study is indicated not only by the variables collected but also by the stratification of the sample. The three categories of subjects, ethnic and gender groupings allow for further understanding of the intricate relationships between MCT status, school behavior and affect, academic performance and school leaving.

References

Beck, M. D. (1977, September). What are pupils' attitudes toward the school curriculum? Elementary School Journal, 73-78.

California State Department of Education, Office of Program Evaluation and Research (1983). A Manual for the Determination of a Severe Discrepancy. Sacramento: State Department of Education.

California State Department of Education (1974). California master plan for special education. Sacramento: Office of State Printing.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart & Winston.

Forness, S. R. (1985). Effects of public policy at the state level: California's impact on MR, LD, and ED categories. Remedial and Special Education, 6(3), 36-43.

Grossman, H J., (Ed.) (1973). Manual on terminology and classification in mental retardation. Washington, D.C.: American Association on Mental Deficiency.

Heber, R. F. (1961). A manual on terminology and classification in mental retardation. (Rev. ed.). American Journal of Mental Deficiency Monograph, (Supp. 64).

Hogan, T. P. (1975). Survey of School Attitudes. New York: Harcourt, Brace, Jovanovich.

Hogan, T. P., personal communication, April 16, 1985 p. 11

Larry P. v. Wilson Riles (1971). Civil Action No. 71-2270 (N.D. Cal., 1971).

Little, T. D., Widaman, K. F., Farren, A. M., MacMillan, A., Hemsley, R. E., & & MacMillan, D.L. (1990a). The factor structure and reliability of the Self Description Questionnaire II (SDQ II) in an early adolescent

popluation, stratified by academic level, ethnicity, and gender. Unpublished technical report (#MCT001). University of California, Riverside, School of Education.

Little, T. D., Widaman, K. F., Farren, A. M., MacMillan, A., Hemsley, R. E., & MacMillan, D. L. (1990b). The factor structure and reliability of the Survey of School Attitudes (SSA) in an early adolescent population, stratified by academic level, ethnicity, and gender. Unpublished technical report (#MCT002). University of California, Riverside, School of Education.

Marsh, H. W. (1986). Self-serving effect (bias?) in academic attributions: Its relation to academic achievements. Journal of Educational Psychology, 78, 190-200.

Marsh, H. W. & Barnes, J. (1982). Self Description Questionnaire II. Sydney, Australia: University of Sydney.

Marsh, H. W., Parker, J., and Barnes, J. (1985). Multidimensional adolescent self-concepts: Their relationship to age, sex, and academic measures. American Educational Research Journal, 22, 422-444.

Nunnally, J. C. (1976). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Public Law 94-142, Education for all handicapped children act, November 29, 1975.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. Review of Educational Research, 46, 407-441.

Slosson, R. L. (1985). Slosson Intellignce Tests for Children and Adults. New York: Slosson Educational Publications.

CHAPTER III

Psychometric Characteristics of the Tests

In determining the impact of minimum competency testing on learning handicapped children, an important consideration is the quality of the minimum competency tests, as defined by their psychometric characteristics. Each cooperating district had constructed its own minimum competency tests, and had also adopted a nationally standardized achievement test series based on district studies which found that the adopted test provided the most valid measure of the instructional objectives of the district. The nationally standardized test administered in each district is used to plot the growth of individual children and also as an instrument of instructional and curricular accountability. In this study, standardized test data were utilized as a criterion for calculating validity, and as the basis for comparison of reliability estimates and difficulty when assessing the quality of the minimum competency tests.

In this chapter of the report the psychometric properties of the proficiency tests used in the four subject districts will be examined. To assess the general quality of the proficiency tests used to determine eligibility for high school graduation, we looked first at their validity, then their reliability, and finally their difficulty.

Validity

The most important psychometric property of a test is its validity, an assessment of what it is that the test measures and how adequately it is measured. The first form of validity, and that which provides the basis for first writing test items is face validity, (a weak form of content validity), the extent to which the test seems to measure the content domain of concern.

The proficiency tests used in this study were constructed by panels of teachers with help from testing experts, and when completed they were analyzed and endorsed by parent/teacher committees and boards of education. From the procedures used in their construction it becomes possible to stipulate that the tests appear to measure the content considered important by these groups.

A second form of validity is criterion-related, the extent to which the test correlates with an independent measure of what the test assesses (Anastasi, 1988). Furthermore, correlations between a new test and an existing test designed to measure the same achievement area provide an estimate of the construct validity of the new test. The reading proficiency tests we examined were designed to assess the students' ability to comprehend text, and the mathematics proficiency tests were designed to assess the students' ability to compute and to solve word problems. For each student in the study, an independent measure of reading comprehension and an independent measure of mathematics ability was available in the form of scores on a nationally standardized achievement test series. For each district, a particular nationally standardized achievement test series had been adopted, as noted above, because it was considered to provide the best assessment of the district's instructional objectives.

The most readily available estimate of the validity of the reading proficiency test was the correlation between score on the proficiency test and score on the total reading test in the standardized achievement test series. Similarly, the most readily available estimate of the validity of the mathematics proficiency test was the correlation between score on the proficiency test and score on the total mathematics test in the standardized achievement test series. These correlations provide an estimate of the amount

of common variance in the two tests, or the extent to which the two measures overlap, and of the extent to which the minimum competency test assesses the reading comprehension or mathematics achievement as measured by the nationally standardized tests. The higher the correlation, the greater the amount of common variance, hence, the greater the commonality of content between the two tests.

In a separate study involving three of the four participating districts, the appropriate minimum competency test scores secured in the spring of the eighth grade year were correlated with the total reading score or the total mathematics score on the nationally standardized test secured at approximately the same time. Those correlations represent the validity coefficients for the total group. However, after the nationally standardized tests were administered to the subjects (Regular Class, Educationally Marginal, and Learning Handicapped) in this study, the minimum competency tests were administered in the normal (i.e., standardized) mode, i.e., in the mode which is used when the test is generally administered to all students. The Pearson Product-Moment correlation coefficient was then calculated on the two sets of scores for the reading tests and for the mathematics tests. Table 3.1 shows the two validity coefficients for the reading and mathematics proficiency tests for the four cooperating districts.

In general, the differences between the validity coefficients for the total group and for the sample selected for this study are not very great. Only District D mathematics shows a difference in validity coefficients that is greater than .1, and all of the coefficients are quite high. Differences would be expected because for the district populations (all eighth graders for districts A, C, and D, and ninth graders for district B), correlations were

based on raw scores; however, for the samples the calculations were based on normal curve equivalent (NCE) scores. Furthermore, each of the samples consisted of a truncated distribution.

Table 3.1

Validity coefficients for the proficiency tests administered in normal mode

	Reading Validity Coefficient		Math Validity Coefficient	
District	$Total_a$	Sample	Total	Sample
A	.657	.579	.773	.748
B	.817	.792	.860	.871
C	.652	.587	.640	.652
D	.758	.799	.776	.627

Note. Correlations with appropriate subtests on the nationally standardized tests were used to determine validity coefficients. $Total_a$ represents the total class population.

The validity coefficients shown in Table 3.1 indicate that both the reading and the mathematics proficiency tests and the standardized tests are measuring substantially common factors, and that the validity of the mathematics minimum competency tests tends to be higher than the validity of the reading minimum competency tests. In general, these coefficients would be considered to be quite high; therefore, it can be suggested that the locally developed reading proficiency tests utilized by these districts are respectably valid instruments for assessing reading comprehension as measured with locally adopted, nationally standardized achievement tests of reading. Similarly, it can be suggested that the locally developed mathematics proficiency tests utilized by these districts are respectably valid instruments for assessing mathematics achievement as measured with nationally

standardized achievement tests of mathematics.

Although the minimum competency tests are respectably valid measures for assessing the competencies that are expected for graduation in these districts, of greater concern in this study is whether the tests are valid for use with the special population of students who were the primary subjects of this study. To determine if the tests administered in the normal mode are as valid for the learning handicapped and the educationally marginal students as for the regular class students, we calculated validity coefficients separately for each group in the manner described above. Table 3.2 shows the validity coefficients for three groups at eighth grade level, combined across districts.

Table 3.2

Minimum competency test validity coefficients across districts by category

Student Category	Reading Validity Coefficient	Mathematics Validity Coefficient
Regular Class	.318	.748
Educationally Marginal	.250	.443
Learning Handicapped	.626	.545

Validity coefficients were expected to be lower for each of the three categories of students because the measurements on each category are truncated. The Regular Class students only include those above the 25th percentile, the Educationally Marginal only include those below the 25th percentile, and the Learning Handicapped tend to be in the lower half of the achievement distribution.

For only one of the four districts was there a large enough sample at the eighth grade level to compare the validity coefficients within the

district for each of the three categories of students, and that was District D. Table 3.3 shows the validity coefficients within that district for each of the three groups.

The reading validity coefficients of Table 3.3 are very similar for each of the groups, but the mathematics validity coefficients are lower for the three groups than for the total population, and they are unexplainably irregular.

Table 3.3

Minimum competency test validity coefficients for District D for three categories of students at grade 8

Student Category	Reading Validity Coefficient	Mathematics Validity Coefficient
Total Grade 8	.758	.776
Regular Class	.754	.288
Educationally Marginal	.752	.588
Learning Handicapped	.705	.642

California's minimum competency law stipulated that school districts were to adopt differential standards for students with diagnosed learning disabilities, and in this study we examined the effects of three alternative modes of test administration -- the consumable mode, where students marked directly on the test booklet rather than a separate answer sheet, the read-out-loud mode, where the examiner read the test material to students while they followed along in the test booklet, and the unlimited time mode, where students were allowed as much time as necessary to complete the examination. For all three modes, test instructions were read out loud to the students.

It was important to learn whether these alternative modes would change

the level of validity of the tests. Tables 3.4 through 3.7 show the validity coefficients for the four modes of administration of the reading and the mathematics proficiency tests for each of the four districts separately.

Table 3.4

Minimum competency test validities for School District A under four administration conditions

Administration Condition	Reading Validity Coefficient	Mathematics Validity Coefficient
Normal mode	.579	.748
Consumable mode	.783	.700
Read-out-loud mode	.701	.838
Unlimited time mode	.651	.838

For the subjects of this study in District A, the consumable mode provided the highest correlation with reading achievement test score, with the read aloud mode next highest. Each alternative mode of administration provided a higher validity coefficient than the normal mode. In mathematics, the read aloud and the unlimited time mode showed the highest correlation with mathematics achievement test score, with the consumable mode the lowest.

Table 3.5

Minimum competency test validities for School District B under four administration conditions.

Administration Condition	Reading Validity Coefficient	Mathematics Validity Coefficient
Normal mode	.792	.871
Consumable mode	.782	.763
Read-out-loud mode	.789	.707
Unlimited time mode	.662	.792

For District B, the normal mode provided the highest validity coefficients for both reading and mathematics. The lowest validity coefficient for reading was in the unlimited time mode, and the lowest for mathematics was the read aloud mode.

Table 3.6

Minimum competency test validities for School District C under four administration conditions

Administration Condition	Reading Validity Coefficient	Mathematics Validity Coefficient
Normal mode	.587	.652
Consumable mode	.854	.760
Read-out-loud mode	.832	.875
Unlimited time mode	.670	.790

For District C, the consumable mode provided the highest validity coefficient for reading and the read aloud mode was highest for mathematics, with normal mode being lowest for both.

Table 3.7

Minimum competency test validities for School District D under four administration conditions

Administration Condition	Reading Validity Coefficient	Mathematics Validity Coefficient
Normal mode	.799	.627
Consumable mode	.718	.806
Read-out-loud mode	.652	.702
Unlimited time mode	.831	.863

In District D, the unlimited time mode provided the highest validity coefficients for both reading and mathematics. The normal mode was next for reading, and the consumable mode was second highest for mathematics.

It is obvious from a perusal of tables 3.4 through 3.7 that there is no one mode of administration that provides the highest validity in all cases. Normal mode provides the highest validity coefficient in one district, unlimited time the highest in one district and tied for highest in math in one district. The consumable mode is highest in reading in two districts, and the read-out-loud mode is highest in mathematics in two districts. The consumable mode is never highest in mathematics, and the read-out-loud mode is never highest in reading. Therefore, it seems reasonable to suggest that alternative modes of administration should be provided because the validity of the measurement will be higher in many instances than when students are restricted to normal mode administration.

Reliability

The reliability of a test is important because it helps to determine the accuracy of decisions made from the test. For example, if the reading proficiency test used in a district has an estimated reliability of 0.70, 70

percent of the variance of the test scores is non-error (Cronbach, 1970) or true score variance, while 30 percent is due to error -- item selection error and test situation error. The larger the amount of error variance in a test, the larger the number of errors that will be made in decisions that are based on the test.

Reliability of the proficiency tests was estimated by two different procedures, depending upon the type of data available to us. In the first situation, reliability had to be estimated from one administration of the test where only total scores for students were available. This included the results from the standardized tests and from the normal mode administration of the proficiency tests.

When item data are not available, and total scores from one administration are used to assess reliability, the Kuder-Richardson Formula 21 (KR21) is utilized. This formula is based upon two fundamental assumptions: that items in one form of a test have as much in common as do the items from parallel forms of the same test, and that items are essentially of equal difficulty. KR21 is a measure of the homogeneity of the test items -- the extent to which each item measures the same thing -- in this case reading comprehension or mathematics performance. Because the error term is limited to variation in the testing situation and variation in the items used, and does not include error associated with the temporal stability of responses, it provides a higher estimate of reliability than that secured using the parallel forms method; however, when item difficulties vary markedly in difficulty, the true reliability coefficient is underestimated.

In the second situation, alternative administrations of the proficiency tests, reliability was estimated using Cronbach's Alpha, which is identical to

Kuder-Richardson Formula 20 (KR20) when items are scored 1 or 0, which is also calculated from one administration of one form of a test. This estimate was possible because we had item data for the target groups on the proficiency tests. This formula takes into consideration all of the data available in individual items, and provides a more accurate assessment of test reliability.

In summary, the reliability of normal mode administration of the proficiency tests and reliability of the standardized achievement tests were estimated using KR21, because we could not secure item data for them. Reliability for the alternative modes of administration of the proficiency tests was estimated using Cronbach's alpha (which, in this situation, is the same as KR20). In this study item data were available for the alternate modes of administration of the minimum competency tests for each of the student categories: regular class (RC), educationally marginal (EM), and learning handicapped (LH), and for each of the modes of administration.

For three of the cooperating districts data were available to us for very large samples of students who had completed the tests as administered in normal mode (Balow, MacMillan, & Hendrick, 1986). For the fourth district the reliability estimate was based on a much smaller sample of students. Table 3.8 shows the KR21 reliability estimates for the proficiency tests and for the standardized tests.

Table 3.8

KR21 reliability estimates for the four cooperating districts on the reading and mathematics proficiency and nationally standardized tests administered in the normal mode

District	Proficiency Test		Standardized Test	
	Reading	Mathematics	Reading	Mathematics
A	.86	.91	.94	.90
B	.89	.86	.92	.93
C	.90	.92	.84	.78
D	.93	.92	.91	.90

The reliabilities reported in Table 8 appear to be quite high for the proficiency tests and for the standardized tests. However, the proficiency test is a "high stakes test" where the passing or failing of the test may have profound consequences on the academic future of the student. Measurement experts are united in believing that when the decision to be made from a test is extremely important and/or irreversible, high reliability is necessary. Truman Kelly (1927) devised reliability standards for tests based on a set of conditions defining the amount of error to be tolerated. He indicated that a reliability of .94 is necessary to evaluate the level of individual accomplishment.

While most of the reliabilities shown in Table 3.8 seem quite close to .94, The Spearman-Brown Prophecy Formula makes it possible to determine how much longer a test would need to be to get from a measured reliability to a higher reliability by using items of equal quality. Using that formula, a test with a reliability of .89 would need to be twice as long to have a reliability of .94; a test with a reliability of .86 would need to be three times as long to attain a reliability of .94; and a test with a reliability of

.78 would need to be almost five times as long to reach that standard. Therefore, the tests in use in the cooperating districts are only marginally reliable in terms of their ability to detect differences as small as one-fourth of a standard deviation.

Grouping into the three alternative modes of test administration followed a stratified random sampling procedure. Subjects were first grouped by category, then within each category they were broken down into ethnic groups and from there into gender groups. This procedure resulted in strata of Regular Class, Anglo, Females; Regular Class, Anglo, Males; Regular Class, Chicano, Females; Regular Class, Chicano, Males; Educationally Marginal, Anglo, Females; Educationally Marginal, Anglo, Males, etc.. From each stratum, subjects were assigned randomly to one of the three test modes.

The minimum competency tests in the cooperating districts were constructed to look very much like achievement tests. They were not devised to discriminate well at the criterion level, and indeed, the criterion level was largely arbitrary. For example, in one district the community-professional advisory committee worked very hard at establishing consensus on the criterion level, and eventually established it at the level of 60 percent correct. However, when this recommendation went to the Board of Education, it decided that 60 percent was too low, and raised the criterion level to 70 percent correct. In practice in three of the four districts, approximately 50 percent of the students fail the proficiency tests the first time they are given.

When tests are designed to discriminate at the mean, it poses problems for decision making with students who are generally inhabiting the tail of the distribution. Therefore, we were concerned with whether the proficiency tests

would be unreliable for our target students, the learning handicapped. To allay this concern, we calculated the test reliabilities for regular class students, the educationally marginal students, and the learning handicapped students on the alternative modes of administration. Tables 3.9 through 3.12 report on those calculated reliabilities on the proficiency tests for each of the four cooperating districts. The reliability coefficients reported are Cronbach's Alpha (also KR20) which gives higher estimates of reliability than KR21. Population reliabilities are KR21 and are taken from Table 3.8.

Table 3.9
Minimum competency test Coefficient Alpha reliabilities across modes of administration for School District A according to three categories of students

Student Category	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.860	.910
Regular Class	.759	.849
Educationally Marginal	.670	.797
Learning Handicapped	.649	.887

According to Table 3.9, the reliability coefficient for each category of student on the alternative modes of administration, are lower than the KR21 estimated reliability for all students under the normal mode administration. Furthermore, the reliability coefficient for the reading proficiency test in District A has inadequate reliability for regular class students and decreases for the marginal students and decreases further for the learning handicapped. Although the mathematics proficiency test is more reliable for each group, it is still of questionable reliability for making significant decisions about individuals. Interestingly, the mathematics test is more reliable for the

learning handicapped than for the other two groups.

Table 3.10

Minimum competency test coefficient alpha reliabilities across modes of administration for School District B according to three categories of students

Student Category	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.890	.860
Regular Class	.647	.723
Educationally Marginal	.831	.811
Learning Handicapped	.680	.854

According to Table 3.10, it is also true for District B that reliabilities for each group combining across all alternative modes of administration are lower than the KR21 estimates for the population on normal mode administration. However, the reliability of the mathematics proficiency test for learning handicapped students taking the alternative modes is almost exactly the same as the KR21 estimate for the total group under normal mode.

Table 3.11

Minimum competency test Coefficient Alpha reliabilities across modes of administration for School District C according to three categories of students

Student Category	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.900	.910
Regular Class	.939	.723
Educationally Marginal	.746	.811
Learning Handicapped	.907	.854

For the reading test administered in alternative modes, the reliability estimate is significantly lower than the KR21 estimate for the population in normal mode only for the Educationally Marginal students. However, the mathematics test reliability estimate is lower than the population estimate in every instance.

Table 3.12

Minimum competency test Coefficient Alpha reliabilities across modes of administration for School District D according to three categories of students

Student Category	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.930	.920
Regular Class	.896	.851
Educationally Marginal	.796	.851
Learning Handicapped	.829	.944

For District D Table 3.12 shows that the estimated reliabilities for the various categories of students taking the reading test under alternative modes are lower than the population KR21 estimate. On the mathematics test, the

learning handicapped students have a higher estimate than the population, the other two groups are lower.

As would be suspected, for each of the three categories of students there is generally a lower reliability coefficient than for the total group. Each of the three categories of students represents a group with restricted range of score (even the Regular Class group is restricted to the upper 75 percent of the student population).

We were also concerned about the reliability of the proficiency tests for the different modes of administration. Tables 3.11 through 3.14 show the estimated reliabilities of the alternative modes of administration across categories of students.

Table 3.13

Minimum competency test Coefficient Alpha reliabilities across categories of students for School District A according to three modes of administration

Mode of Administration	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.860	.910
Consumable	.707	.894
Read-Out-Loud	.756	.870
Unlimited Time	.675	.924

None of the alternative modes of administration result in reliability estimates for the reading test which are acceptable for making significant decisions about individuals. The mathematics test proves to be more reliable, but is marginal at best for making individual decisions. Only for the unlimited time mode of administration of the mathematics test is the reliability as high as for the normal mode administration for the school population.

Table 3.14

Minimum competency test Coefficient Alpha reliabilities across categories of students for School District B according to three modes of administration

Mode of Administration	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.890	.860
Consumable	.893	.908
Read-Out-Loud	.850	.869
Unlimited Time	.903	.897

For District B, in most instances the alternative modes of administration reliability estimates are as high or higher than for the normal mode administration reliabilities. Nevertheless, the reported reliability estimates are in the marginal area for decision making about individuals.

Table 3.15

Minimum competency test Coefficient Alpha reliabilities across categories of students for School District C according to three modes of administration

Mode of Administration	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.900	.920
Consumable	.943	.916
Read-Out-Loud	.928	.913
Unlimited Time	.932	.938

In District C, for both tests and all alternative modes of administration, the estimated reliabilities are as high or higher than for the normal mode of administration.

Table 3.16

Minimum competency test Coefficient Alpha reliabilities across categories of students for School District D according to three modes of administration

Mode of Administration	Reading Reliability Coefficient	Mathematics Reliability Coefficient
Population	.930	.920
Consumable	.926	.941
Read-Out-Loud	.885	.904
Unlimited Time	.929	.958

In District D, estimated reliabilities for the alternative modes of administration tend to be about as high as for normal mode administration, with the exception of the read-out-loud mode for the reading test where the reliability is significantly lower. The consumable mode and the unlimited time mode on the mathematics test result in reliabilities clearly high enough for individual decision making.

Difficulty

A third psychometric concern, although much less important in this study, is the difficulty of the proficiency examinations. Because students may be precluded from graduating because of failing the proficiency test developed in the district, the difficulty of the test is quite important. With a very difficult proficiency test only very capable students will pass and be allowed to graduate. With a very easy proficiency test, virtually everyone will pass and be allowed to graduate. Difficult tests will result in a high proportion of students enrolling in remedial classes; easy tests will result in very few students enrolling in such classes.

Because the educationally marginal students in this study and the learning handicapped students in general tend to achieve at a lower level than

other students, easy proficiency tests are advantageous to them. That is, the lower the level of minimum competency adopted by the district, the more likely it is that the educationally marginal and learning handicapped students will be allowed to graduate through completion of normal requirements.

To the extent that minimum competency standards differ by district, parents can shop for districts with the lowest standards in order to enhance the probability of graduation for their educationally marginal or learning handicapped students.

To determine the comparative difficulty of the various tests, a regression analysis was carried out on the set of proficiency and achievement scores for each school district to derive a formula for predicting proficiency test scores from standardized test scores. Using that formula we then calculated the standardized reading score a student would need in order to score at the criterion level on the proficiency test. Substantial variation was found in reading and mathematics achievement levels corresponding to the proficiency test criteria for the cooperating districts. Table 3.17 illustrates this variability based on normal mode population analyses.

Table 3.17

Grade equivalent score on the district adopted standardized achievement test corresponding to criterion score on district minimum competency test for reading and mathematics based on population values

	Minimum Competency Test Subtest	
District	Reading	Mathematics
A	5.1	7.6
B	6.2	7.5
C	7.0	8.8
D	8.3	8.6

For the four districts cooperating in the study there is more than a three year range among them in their expectations for minimum competency in reading, and greater than a one year range in their expectations for mathematics competency.

Discussion

In the four districts included in this study, the validity of the minimum competency tests is quite high when the district adopted achievement test series is used as the criterion. The districts have managed to tap the same content with their locally constructed tests. Furthermore, the tests appear to be quite valid for the learning handicapped students in the samples. Nevertheless, for some children in all categories, mode of administration appears to be an important variable in the validity of the minimum competency tests. No one mode of administration is consistently best in terms of level of validity, but each mode of administration is best with some tests in some district. Inasmuch as allowing students to mark on the test booklet instead of a separate answer sheet (consumable mode), reading the test material to the students (read-out-loud mode), and allowing unlimited time to complete the

test (unlimited time mode) increases test validity, and inasmuch as validity is the most important characteristic of a test, it would appear wise to make these adjustments available to all children.

With respect to reliability, in three of the districts the minimum competency test was substantially lower for the learning handicapped pupils than for the population as a whole. That is, MCT decisions were much less reliable for the learning handicapped pupils than for other pupils in the district. However, mode of administration did not adversely affect the reliability of the tests. In general, regardless of the mode of administration the estimated reliability of the tests tended to remain close to the level found for the grade level population of students. This latter finding also argues in favor of alternative modes of administration for everyone: validity is enhanced and reliability does not go down.

It was found that there were significant differences for each of the districts in the level of difficulty of the reading test as compared with the mathematics test. Passing the MCT in reading in District A required a reading ability of only 5.1 whereas passing the mathematics MCT required a math score of 7.6. The difference in District C was 1.8 years and in District D was .3 years. Differences between districts were just as pronounced.

There was a 3.2 year range in grade equivalent score in reading required for passage of the reading MCT among the four districts, and a 1.2 year range in mathematics scores.

NOTES

According to Kaplan and Saccuzzo (1989) "it is rare to see a validity coefficient larger than .60, and validity coefficients in the range of .30 to .40 are commonly considered high."

References

Anastasi, A. (1988). Psychological Testing, (6th ed.). NewYork: Macmillan.

Balow, I. H., MacMillan, D. L. & Hendrick, I. G. (1986). Local option competency testing: Psychometric issues with mildly handicapped and educationally marginal students, Learning Disabilities Research. 2("1"): 32-37.

Cronbach, L. J. (1970). Essentials of psychological testing. New York: Harper and Row.

Kaplan, R. K. & Saccuzzo, D. P. (1989). Psychological testing: Principles, applications, and issues, (2nd. ed.). Pacific Grove, CA: Brooks/Cole.

Kelly, T. (1927). Interpretation of educational measurements. Yonkers, NY: World Book.

Table 4.28

Results of Univariate Analyses on Math Minimum Competency Test

	$df(df_n, df_d)$	F	p
Category	2, 5	21.51	0.0001
RC vs EM & LH (D1)	1, 5	40.27	0.0001
EM vs LH (D2)	1, 5	9.22	0.0026

Note. Only those effects significant at the .05 level or greater are reported.

Summary

The primary purpose of this study was to determine the effect failure of MCTs has on students of varying academic ability. Specifically, the intent of this study was to examine the Minimum Competency Tests'(MCT) impact on students' educational careers. While the results of so many analyses are quite detailed and interesting, an overall sketch is helpful.

As would be expected, academic ability, as described by category, is predictive of math and reading achievement test performance. As expected, RCs have the highest mean achievement test scores followed by EM and LH scores respectively. Similarly, MCT reading and math scores fell in descending order for the RC, EM, and LH categories.

At this point, the reader should be reminded of the caveats of this study presented previously in this chapter. Primarily, the concern is the intertwining of academic success and academic ability. Ascertaining which comes first, which impacts which, etc., is impossible given the design of this study and the nature of the constructs involved. Nonetheless, the study revealed findings indicating performance on the MCT to be significant in the areas of courses selected, involvement in school, and school affect. In general, the consequences of failure on the MCT are related with several

CHAPTER IV

Consequences of MCT Performance on Student Careers

One purpose of this project was to study the impact of Minimum Competency Tests (MCT) on the educational careers of secondary students. To this end, the educational experiences of those students passing the MCT were compared with the educational experiences of those students not passing the MCT. Some specific outcomes anticipated to be affected by MCT performance were the types of courses taken, transferring out of the district (to a district with an "easier" test), and early school dropout. Other outcomes possibly related to MCT performance such as school attendance, disciplinary problems, self-concept, and attitudes toward school were also examined. Results of the corresponding analyses employed to achieve these aims are reported in this chapter.

An acknowledged caveat of this study is the intertwining of academic success and academic ability. Surely, factors contributing to MCT failure permeate the student's academic life, thus confounding the cause of MCT failure. A multitude of events prior to any failure cannot be disentangled into any single cause. As such, isolating an outcome, such as MCT failure, and then monitoring consequent behavior does not permit an assumption of causal effects. Nevertheless, as in all social science studies, this research necessarily isolates events and consequences.

All analyses were performed on data gathered through the fourth year of the study. This decision was made in consideration of maximizing the completeness of each variable. At the end of the fourth year of this study, cohort one subjects completed 11th grade and cohort two subjects completed

10th grade.

Analysis of Variance (ANOVA) and Multivariate Analysis of Variance (MANOVA) were the statistical methods of choice. When estimating the sums of squares for main and interaction effects, the analysis was performed in a hierarchical manner. Specifically, each main effect was estimated while partialling out the effects of other main effects, each two-way interaction had main effects and other two-way interactions partialled out, and the three-way interactions had all main and two-way interaction effects partialled out (cf. Humphreys & Fleishman, 1976). Significant multivariate effects were followed up by investigating univariate results. When relevant, a multiple comparison was done to determine significant mean differences within a group.

Course Selection

MCT (passed or failed) status, category (LH, EM, RC) and ethnicity were used to predict CLASS, a variable representing the level of difficulty of courses taken by students. After consulting with school personnel, all courses were categorized as remedial courses, required courses for graduation, or elective courses. Remedial, required, and elective courses were assigned a numeric value (1, 2, and 3, respectively) then averaged to create the dependent variable, CLASS, for each student. The results of this analysis are presented in Table 4.1.

Table 4.1

Results of ANOVA on Type of Courses: Required, Remedial & Elective

	df	F	p
Category	2, 917	163.07	0.0001
MCT Status	1, 917	15.31	0.0001
Ethnic	1, 917	0.37	0.6919
Category x MCT Status	2, 909	1.20	0.3017
Category x Ethnic	4, 909	1.70	0.1481
MCT Status x Ethnic	2, 909	0.09	0.9176
Category x Ethnic x MCT Status	4, 905	0.25	0.9114

Note: Observations with missing values were not used in this analysis, thus only 923 observations were used in this analysis.

MCT Status Main Effect

Across Categories, course selection revealed a main effect for MCT status ($F(1, 917) = 15.31$, $p < .0001$). As hypothesized, those students not passing the MCT, regardless of category, had a lower mean value for CLASS ($M = 1.92$) than those passing the MCT ($M = 2.10$). This indicates students who failed the MCT were enrolled in more remedial and required courses than students who had passed the MCT. Because MCT performance was a two level variable (pass/fail) a multiple comparison test was irrelevant; a significant prediction by a two-level variable implicity indicates a difference between the two levels of the variable.

Category Main Effect

Category was also a predictor of course selection, $F(2, 917) = 163.07$, $p < .0001$. This univariate analysis was followed up with a multiple comparison

to determine any significant mean differences between study groups. A Bonferonni multiple comparison revealed differences between all three study groups. Means followed the expected trend, with students in the LH sample ($\underline{M}$ = 1.72) enrolled in more remedial courses than the EM ($\underline{M}$ = 1.99) and RC ($\underline{M}$ = 2.15) students.

Ethnic Main Effect

Ethnicity was not a significant predictor of courses selected. There were no mean differences between ethnic groups in the type of courses selected.

Interactions

Although both MCT status and category were main effects, the interaction between them was not significant in predicting courses selected. This suggests that there was not a difference in courses selected between those who failed versus passed the MCT among Categories.

Summary of Course Selection

Course selection differs depending on a student having passed or failed the MCT. This difference in the pattern of courses taken is present regardless of Category. Students passing the MCT enroll in more elective courses and fewer remedial courses than do students failing the MCT.

As expected, the Category variable also predicted the type of courses in which students enrolled. The pattern of means indicated that the LH group enroll in more remedial courses than did the EM group, and the RC students enrolled in the fewest remedial courses.

Course selection did not differ between ethnic groups; nor were differences in the interactions among the ethnicity, Category and MCT status found.

Clearly performance on MCTs implicitly impacts courses selected. The MCT could be considered as a "gate" which opens for students passing MCTs and closes for those who fail. This gating role of the MCT allows those students passing the MCT the greatest selection of courses from which to choose, as their course schedule consists primarily of required or remedial courses. Both the IEP and differential standards provide a back-door entrance to elective courses for the more protected LH student.

School Involvement

A multivariate analysis (MANOVA) was performed to determine significant predictors of School Involvement and dropout. School Involvement is a composite variable consisting of attendance behavior and discipline referrals. Dropout is a dichotomous variable, 1 represents a dropout, 0 represents persisting. Results of this analysis are outlined in Table 4.2. Based on the MANOVA results, separate univariate analyses were completed for the computed on the school involvement and the dropout variables.

Attendance information was gathered from students' cumulative files as well as from printouts generated by the school's attendance office and districts' data processing. The number of days per year that a student missed school over the five years of study were averaged to create the attendance variable. This is a continuous variable ranging from 0 to 83.5 days absent.

A combination of the various types of discipline referrals (e.g., suspension, anti-teacher, student-conflict, maladjusted behavior, substance abuse and unspecified) was used to create a comprehensive discipline variable for each grade level. The discipline-by-grade variable indicates when a subject received one referral (assigned value of 1), 2 referrals (assigned value of 2), and 3 or more referrals (assigned value of 3) for any combination

of the above behaviors. If the student received no referrals, the value of the discipline variable was zero. The average discipline value was taken for all grade levels to create a discipline variable representative of the students' high-school career. These discipline data were gathered at the school site from either the student cumulative record files or discipline records, whichever yielded more information.

Attrition data included information on dropout and school transfer. As of June 1990, all 1188 subjects were accounted for as being still enrolled in one of the schools in the study, a dropout, or a transfer. These analyses focus on the in-school and dropout classifications. Classification of transfer indicates that student is no longer in the original district or comprehensive high-school with the rest of their peers. A transfer student was classified as such only if the school could provide information indicating that there had been a request for transcripts, or the new school's address. A student was classified as a dropout when school records revealed any of the following codes: No show; Expulsion; Withdrawn; Juvenile Hall; Runaway; Overage; Job Corp; and Unknown. The information and ID number of one subject who died during the study was removed from the data set.

Table 4.2

Results of MANOVA on School Involvement & Dropout

	$df(df_n, df_d)$	F	p
Category	6, 1114	2.58	0.0173
MCT Status	3, 557	9.16	0.0001
Ethnic	6, 1114	1.79	0.0969
Category x Ethnic	12, 1453	2.00	0.0213
Category x MCT Status	6, 1098	0.93	0.4739
Ethnic x MCT Status	6, 1098	0.66	0.6813
Category x Ethnic x MCT Status	12, 1442	0.67	0.7841

Note: Observations with missing values were not included in the analyses, thus, only 565 observations were used.

MCT Status Main Effect

Of the three factors, (MCT status, category, and ethnicity) only the main effect of MCT status was significant, $F(3,557) = 9.16$, $p < .0001$, although category did approach significance. Table 4.3 presents a further look at the univariate analyses and reveals a difference in attendance patterns between those failing and passing MCTs ($F(1,559) = 17.82$, $p < .0001$). The mean number of days of school missed for those passing the MCT was 11.65 and for those failing the MCT the mean days absent was quite a bit higher, at 19.52

Table 4.3

Results of ANOVA on Attendance

	df	F	p
Category	2, 559	4.49	0.0117
Ethnic	2, 559	2.36	0.0957
MCT Status	1, 559	17.82	0.0001

Table 4.4

Results of ANOVA on Dropout

	df	F	p
Category	2, 559	0.15	0.8630
Ethnic	2, 559	0.07	0.9370
MCT Status	1, 559	11.30	0.0008

The analysis presented in Table 4.4 considers the impact of MCT status on the student's decision to continue in, or drop out of, school. The dropout variable was coded "0", for in school or transferred to another school, and "1" for dropout behavior. MCT status had a main effect on dropout ($F(1,559) = 11.30$, $p < .001$), with students failing the MCT having a higher mean dropout rating ($M = 0.17$) than those who passed the MCT ($M = .08$). A separate table for the univariate analysis of discipline referrals is not presented here as MCT status was not related to discipline referrals, ($p > .10$).

Table 4.5

Means and Standard Deviations of School Involvement Variables, by MCT Status, Category and Ethnicity

	Regular Class				Educationally Marginal				Learning Handicapped			
	Passed		Failed		Passed		Failed		Passed		Failed	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
						Class						
A'	2.16	.26	2.12	.24	2.04	.16	1.94	.22	1.82	.31	1.70	.26
B'	2.17	.21	2.11	.19	2.11	.11	2.07	.16	1.82	.28	1.65	.30
C'	2.16	.19	2.12	.17	2.01	.20	1.93	.22	1.78	.22	1.72	.29
						Attendance						
A'	10.13	10.30	16.58	17.49	11.75	8.75	17.93	13.29	7.92	6.16	21.16	18.93
B'	9.22	10.61	12.27	17.47	24.86	25.79	27.11	20.70	6.50	4.94	16.75	19.09
C'	13.23	11.97	15.26	11.30	16.96	15.57	24.56	24.60	16.80	13.75	2'.25	27.26
						Discipline						
A'	.10	.25	.15	.27	.24	.34	.25	.42	.17	.32	.27	.43
B'	.18	.28	.19	.32	.39	.66	.45	.67	.25	.35	.57	.75
C'	.15	.37	.25	.38	.22	.34	.18	.36	.00	.00	.45	.65
						Drop out						
A'	.62	.25	.19	.39	.08	.26	.22	.42	.00	.00	.15	.35
B'	.08	.27	.27	.46	.00	.00	.00	.00	.00	.00	.38	.52
C'	.10	.31	.15	.36	.09	.29	.14	.34	.20	.44	.15	.36

Note. A'= Anglo; B' = Black; C' = Chicano.

Summary of School Involvement

Clearly, MCT status impacts on educational experiences. The school involvement variable encompasses major dimensions of a students' experience, specifically attendance and dropout behavior. That MCT status is a predictor of these behaviors indicates the relevance of MCT performance on students' educational careers.

It was somewhat surprising that category and ethnicity did not effect the school involvement variable. As such, regardless of a students' academic achievement and ethnicity, MCT performance is the best predictor of school involvement. Again, the caveat presented at the beginning of this chapter should be considered here. Are students who the fail the MCT less "involved" (as defined by our school involvement variable) because they failed the MCT or do they fail the MCT because they are not "involved" in school?

School Affect

Academic and non-academic self-concept, and attitudes toward school were considered as measures of school affect. Three separate MANOVAs were employed to determine how the predictor variables (MCT status, category and ethnicity) contributed to differences in school affect.

Academic Self-Concept Scales

The MANOVA on academically-related self-concept scores used four self-concept scales as dependent variables: General, Academic, Verbal, and Math self-concept.

Table 4.6

Results of MANOVA on Four Academic Self-Concept Scales

	$df(df_n, df_d)$	F	p
Category	8, 2296	29.044	0.0001
MCT Status	4, 1148	8.6595	0.0001
Ethnic	8, 2296	6.4311	0.0001
Category x MCT Status	8, 2280	1.5391	0.1385
Category x Ethnic	16, 3483	2.058	0.0077
MCT Status x Ethnic	8, 2280	0.5246	0.8388
Category x MCT	16, 3471	1.944	0.0134

MCT status main effect. The multivariate test of differences between MCT status on Academic self-c ncept was statistically significant, $F(4,1148) = 8.66$, $p < .001$. Based on the univariate results, MCT status was significant in predicting the General, Academic and Math sub-scales of Academic self-concept (all ps < .05). However, Verbal self-concept was not significantly related to MCT status ($p > .05$). See Table 4.6 for figures.

Category main effect. The multivariate test of differences between the three categories of students on these variables was also statistically significant, $F(8,2296) = 29.04$, $p < .0$. Univariate results are presented in Table 4.7; RC, EM and LH students differed significantly on each of these four dimensions of self-concept (all ps < .001). Mean values for categories were in the predicted direction, with the RCs having the highest mean on all four dimensions, EMs slightly lower and LHs with the lowest. The exception to this was on General self-concept where EMs and LHs were virtually equal (EM M = 4.78, LH M = 4.79). See Table 4.9 for means and standard deviations.

Ethnic main effect. The multivariate test of differences between the three ethnic groups on academically-related self-concept scales was statistically significant, $F(8,2296) = 6.43$, $p < .0001$ (Table 4.6).

Univariate analysis (see Table 4.7) demonstrate that Math was the only self-concept sub-scale on which the ethnic groups did not differ significantly ($p > .10$). The remaining three academically related self-concept scales showed significant group differences between the three ethnic groups (all $ps < .001$). Interestingly, means failed to represent the predicted patterns of differences among the three ethnic groups. Somewhat unexpectedly, Black students had the highest mean scores on each of the four academic dimensions of self-concept, and Anglo and Chicano students had lower, and approximately equal, levels of self-concept on all four scales. These trends were supported by Dunn's test results, which showed that Anglo and Chicano students differed nonsignificantly on all four of the scales (all $ps > .20$), and that Black students had higher self-concept scores than did Anglo and Chicano students on all four dimensions of academically-related self-concept (all $ps < .01$). Group means are presented in Table 4.7.

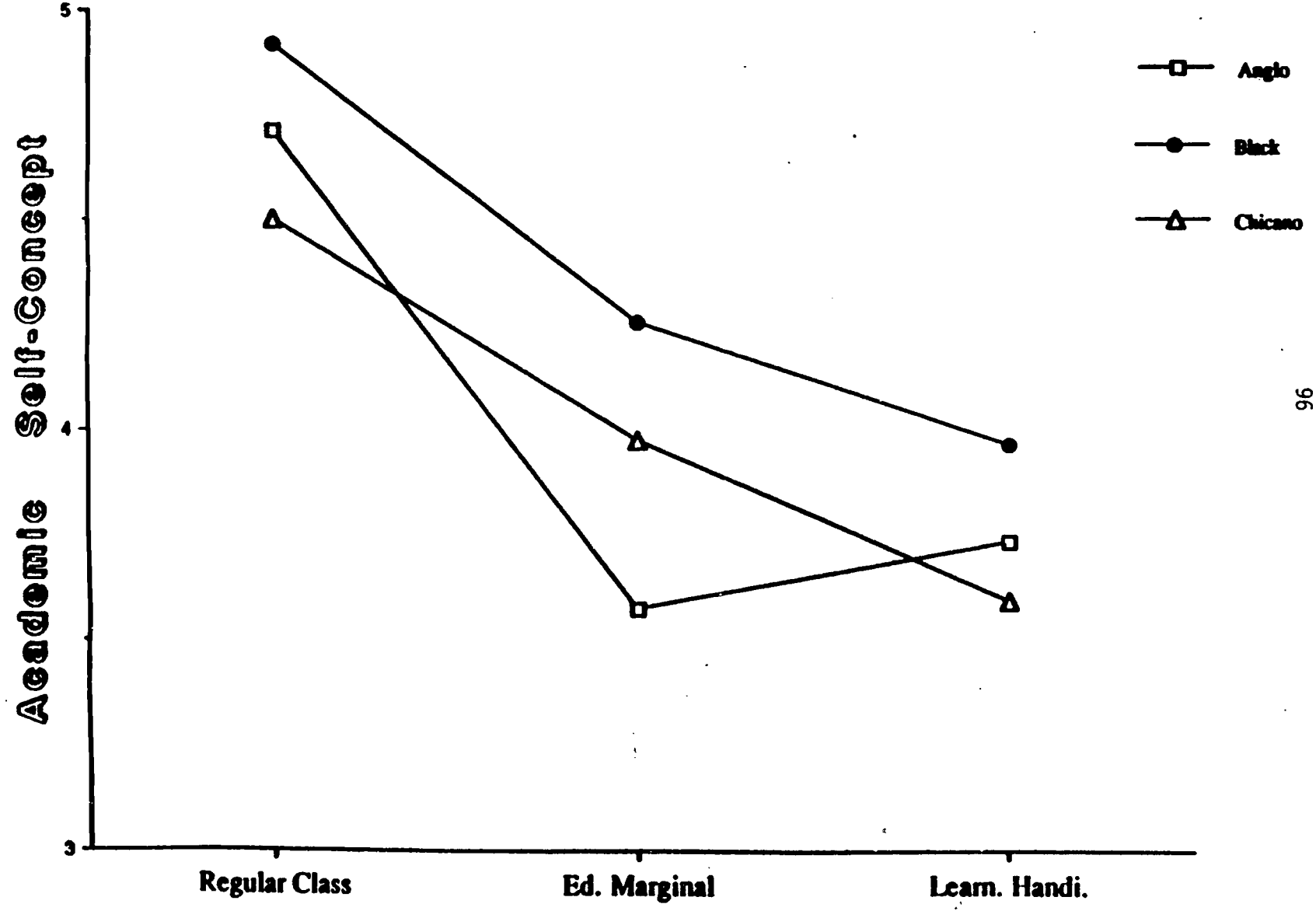
Interactions. The multivariate test of the category X ethnicity interaction on academically-related self-concept was the only statistically significant multivariate interaction, $F(16,3483) = 2.06$, $p < .01$ (see Table 4.6).

Only one of the four univariate interactions was significant. This was on the Academic self-concept scale, $F(4,1143) = 3.49$, $p < .01$. As shown in Figure 4.1, the trends in Academic self-concept for the Black and Chicano students parallel their levels of achievement, RC students having the highest levels of Academic self-concept, LH students the lowest levels, and EM

students scoring about midway between the RC and LH students. A similar pattern was found for RC and LH Anglo students, with RC students scoring at much higher levels than LH students. But, Anglo EM students scored at much lower levels than would be predicted from their levels of academic achievement, scoring below Anglo LH students, rather than midway between RC and LH students. In fact, the significant interaction arose solely from the decidedly low scores for the Anglo EM students.

Figure 4.1.

Academic self-concept, by Achievement and Ethnic group



The remaining two two-way interaction effects (category X MCT status and MCT status X ethnicity) for the academically related self-concept scales were nonsignificant as was the test of the three-way interaction (all ps > .10). As a result, univariate results for these interaction effects were not investigated further.

Table 4.7

Results of Univariate Analyses on Academic Self-Concept Scales

	df	F	p
	General Self-Concept		
Category	2, 1151	17.11	0.0001
MCT Status	1, 1151	6.00	0.0145
Ethnic	2, 1151	12.53	0.0001
Category x MCT Status	2, 1143	0.73	0.4803
Category x Ethnic	4, 1143	1.56	0.1819
MCT Status x Ethnic	2, 1143	0.74	0.4754
Category x MCT Status x Ethnic	4, 1139	3.23	0.0121
	Academic Self-Concept		
Category	2, 1151	82.57	0.0001
MCT Status	1, 1151	22.73	0.0001
Ethnic	2, 1151	11.09	0.0001
Category x MCT Status	2, 1143	0.47	0.6231
Category x Ethnic	4, 1143	3.49	0.007
MCT Status x Ethnic	2, 1143	1.54	0.2142
Category x MCT Status x Ethnic	4, 1139	2.27	0.0595

	Verbal Self-Concept		
Category	2, 1151	111.41	0.0001
MCT Status	1, 1151	3.24	0.0001
Ethnic	2, 1151	12.41	0.0001
Category x MCT Status	2, 1143	1.16	0.3132
Category x Ethnic	4, 1143	1.45	0.2168
MCT Status x Ethnic	2, 1143	0.66	0.5195
Category x MCT Status x Ethnic	4, 1139	1.76	0.1341
	Math Self-Concept		
Category	2, 1151	10.25	0.0001
MCT Status	1, 1151	29.98	0.0001
Ethnic	2, 1151	0.89	0.4123
Category x MCT Status	2, 1143	1.46	0.2329
Category x Ethnic	4, 1143	2.17	0.0707
MCT Status x Ethnic	2, 1143	0.99	0.3720
Category x MCT Status x Ethnic	4, 1139	2.39	0.0491

Summary of Academic Self-Concept. In general, students who failed MCT differ on academic self-concept from those who passed. Specifically, those students who passed the MCT displayed significantly more favorable General, Academic and Math self-concepts in comparison with students who failed MCT. However, they did not differ from those failing MCT on Verbal self-concept.

The MCT status variable was predictive of academic self-concept, which could indicate that MCT is proxy for a number of indicators of the academic success in a student's life. As indicated by the interaction between ethnicity

and category, the academic level at which students perform effect the academic self-concept of ethnic groups differentially. Specifically, Black students exhibited the highest academic self-concept across category. This finding was quite suprising, given the mean academic achievement of Black students was not the highest in any category (see table 4.21 for means of standardized achievement scores).

One could speculate that the higher self-concept of Black subjects resulted because the sensitization of school personnel to the need to encourage minority students encountering academic difficulties. The implicit belief of teachers could be that white students receive needed supports at home. As self-concept can be impacted by the responses received from significant others, these minority students may be receiving messages from well-meaning teachers that their academic achievement is satisfactory. Apparently, white students with academic problems (i.e., EM) are personalizing failure experiences, as the self-concepts of white EM subjects were lower than academic self-concept scores of white LH subjects.

Nonacademic Self-Concept Scales

The MANOVA on nonacademically-related self-concept scores used the remaining seven self-concept scales as dependent variables, specifically the Honesty, Emotions, Parents, Physical, Appearance, Same Sex relations, and Opposite Sex relations scales.

Table 4.8

Results of MANOVA on Seven Non-Academic Self-Concept Scale

	$df(df_n, df_d)$	F	p
Category	14, 2290	4.888	0.0001
MCT Status	7, 1145	4.5608	0.0001
Ethnic	14, 2290	4.8189	0.0001
Category x MCT Status	14, 2274	0.7371	0.7382
Category x Ethnic	28, 4101	1.0098	0.4503
MCT Status x Ethnic	14, 2274	0.7966	0.6739
Category x MCT Status x Ethnic	28, 4086	0.8546	0.6850

Category main effect. The multivariate test of differences between the three categories of students on these variables was statistically significant, $F(14,2290) = 4.89$, $p < .0001$. Univariate results indicate category to be significant predictor of Honesty, Appearance and Same Sex relations subscales (see Table 4.11).

Results of multiple comparison tests indicate RC students had significantly higher scores than did the EM and LH students on five of the seven scales (i.e., Honesty, Emotions, Physical, Appearance, and Opposite Sex relations), and the EM and LH students differed nonsignificantly on all six of the scales. On Same Sex relations, EM students had a higher level of self-concept than did the RC students (Dunn's test; $p < .05$), and RC and LH students differed significantly ($p > .05$).

MCT Status. The multivariate test of difference between the two levels of MCT status on the non-academic self-concept variables was significant,

$F(7,1145) = 4.56$, $p < .0001$ (see Table 4.8).

Examination of the univariate results indicate that Honesty is the only scale significantly predicted by MCT status ($F(1,1151) = 15.03$, $p < .0001$). At the .05 level of significance, however, Same Sex relations is predicted by MCT status. Univariate results are summarized in Table 4.11.

Ethnic main effect. The main effect of ethnicity on the nonacademic self-concept scales was significant ($F(14,2290) = 4.82$, $p < .0001$) (see Table 4.8). Group means are presented in Table 4.10. As a main effect, ethnicity was significant for Honesty, Physical Ability, Appearance, and Opposite Sex scales. For the Physical Ability scale, ethnicity was the only significant predictor.

Dunn's multiple comparison test of differences showed a ($p < .05$) difference between Anglos and Chicanos on the Honesty self-concept scale, with Anglos having the higher score. Blacks and Chicanos differed significantly on Emotional self-concept, with Chicanos scoring lower than Blacks (see Table 4.10). Significant differences between Blacks and Anglos, and Blacks and Chicanos were found on the Physical (Black $M = 5.01$, Anglo $M = 4.67$, Chicano $M = 4.63$), Appearance (Black $M = 4.55$, Anglo $M = 4.06$, Chicano $M = 3.94$), and Opposite Sex (Black $M = 4.90$, Anglo $M = 4.55$, Chicano $M = 4.50$), scales. Blacks and Anglos differed on Same Sex scale (Black $M = 4.97$, Anglo $M = 4.81$), and no significant differences were found between ethnic groups on the Parent scale.

Interactions. No two- or three-way interactions were significant on the multivariate level (see Table 4.8). As such, neither multivariate nor univariate interactions are discussed.

Table 4.9

Means and Standard Deviations of Academic Self-Concept Variables, by MCT Status, Category and Ethnicity

	Regular Class				Educationally Marginal				Learning Handicapped			
	Passed		Failed		Passed		Failed		Passed		Failed	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
						General						
A*	5.14	.59	4.98	.76	4.81	.73	4.79	.79	5.05	.69	4.83	.75
B*	5.41	.41	5.27	.54	5.11	.76	4.97	.97	3.67	1.36	4.96	.77
C*	5.07	.65	4.91	.76	4.74	.69	4.65	.88	4.87	.50	4.56	.72
						Academic						
A*	4.73	.77	4.55	.83	3.97	.77	3.53	1.07	4.33	.92	3.68	.96
B*	4.93	.67	4.82	.77	4.37	.64	4.28	.88	3.18	1.12	3.95	.83
C*	4.62	.70	4.29	.96	4.12	.78	3.80	1.01	3.61	.62	3.71	.97
						Verbal						
A*	4.37	.83	4.38	.81	3.59	.73	3.42	.93	3.92	.89	3.37	.88
B*	4.56	.70	4.56	.92	3.88	.66	3.96	.86	3.07	.59	3.52	.97
C*	4.27	.83	4.02	.87	3.40	.86	3.46	.94	3.33	.81	3.23	.85
						Math						
A*	4.18	1.10	3.76	1.15	3.73	1.23	3.15	1.35	3.74	1.44	3.28	1.27
B*	4.17	1.17	3.83	1.37	4.01	.71	3.43	1.25	1.71	.69	3.66	1.41
C*	3.93	1.16	3.40	1.33	3.84	1.20	3.28	1.18	3.77	1.59	3.62	1.25

Note. A* = Anglo; B* = Black; C* = Chicano.

Table 4.10

Means and Standard Deviations of Non-Academic Self-Concept Variables, by MCT Status, Category and Ethnicity

	Regular Class				Educationally Marginal				Learning Handicapped			
	Passed		Failed		Passed		Failed		Passed		Failed	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
						Honesty						
A*	4.58	.81	4.32	.83	4.33	.73	4.17	.96	4.70	.88	4.15	.87
B*	4.41	.72	4.13	.85	3.96	.72	4.08	.90	3.45	1.32	4.11	.65
C*	4.41	.64	4.28	.90	4.13	.84	3.97	.97	4.20	.71	3.96	.87
						Emotion						
A*	4.30	.80	4.14	.84	4.00	.84	4.07	.87	4.45	.79	4.13	.84
B*	4.43	.76	4.40	.96	4.47	.94	4.17	1.09	3.04	1.05	4.07	.84
C*	4.23	.80	4.11	.86	4.09	.73	3.98	.88	4.32	.81	4.07	.76
						Parent						
A*	4.60	1.06	4.37	1.18	4.26	1.20	4.30	1.13	4.60	1.00	4.45	.91
B*	4.55	1.01	4.65	1.17	4.42	1.07	4.57	1.04	3.18	.90	4.55	.97
C*	4.52	1.03	4.47	1.02	4.42	.98	4.40	1.00	4.32	1.35	4.57	.83
						Physical						
A*	4.76	.90	4.72	1.05	4074	1.06	4.65	.92	4.91	.87	4.46	.98
B*	5.11	.63	5.13	.68	5.17	.94	4.76	.89	4.93	.58	4.74	.90
C*	4.76	.88	4.68	.88	4.58	.89	4.53	.98	4.60	.87	4.56	.82
						Appearance						
A*	4.14	.88	4.17	1.04	3.84	1.05	3.88	1.08	4.26	.71	4.12	1.07
B*	4.57	.73	4.64	.69	4.22	1.19	4.58	1.20	3.65	1.68	4.50	1.06
C*	4.06	.93	4.21	.96	3.75	1.04	3.81	1.06	3.59	1.14	3.84	1.03
						Same Sex						
A*	5.01	.62	4.83	.80	4.70	.79	4.76	.83	4.73	.82	4.64	.72
B*	5.16	.52	5.09	.66	5.00	.52	4.82	.85	4.15	1.24	4.46	.88
C*	5.09	.56	4.97	.68	4.84	.64	4.63	.79	4.80	.41	4.48	.71
						Opposite Sex						
A*	4.59	.82	4.55	1.01	4.47	.98	4.56	.91	4.50	.98	4.53	.96
B*	4.89	.77	5.04	.86	4.87	.62	4.97	.90	3.74	1.96	4.76	.87
C*	4.64	.72	4.73	.92	4.35	.96	4 38	.92	4.28	.91	4.34	.81

Note. A* = Anglo; B* = Black; C* = Chicano.

Table 4.11

Results of Univariate Analyses on Non-Academic Self-Concept Scales

	df	F	p
	Honesty Self-Concept		
Category	2, 1151	7.05	0.0009
MCT Status	1, 1151	15.03	0.0001
Ethnic	2, 1151	5.68	0.0035
Category x MCT Status	2, 1143	0.61	0.5451
Category x Ethnic	4, 1143	0.34	0.8488
MCT Status x Ethnic	2, 1143	0.68	0.5066
Category x MCT Status x Ethnic	4, 1139	1.53	0.1900
	Emotion Self-Concept		
Category	2, 1151	3.79	0.0229
MCT Status	1, 1151	2.94	0.0865
Ethnic	2, 1151	2.50	0.0825
Category x MCT Status	2, 1143	0.43	0.6523
Category x Ethnic	4, 1143	1.62	0.1658
MCT Status x Ethnic	2, 1143	0.55	0.5799
Category x MCT Status x Ethnic	4, 1139	2.09	0.0796

	Parent Self-Concept		
Category	2, 1151	2.30	0.1007
MCT Status	1, 1151	0.27	0.6067
Ethnic	2, 1151	0.69	0.5032
Category x MCT Status	2, 1143	1.20	0.3024
Category x Ethnic	4, 1143	0.69	0.6001
MCT Status x Ethnic	2, 1143	2.11	0.1220
Category x MCT Status x Ethnic	4, 1139	1.11	0.3501
	Physical Self-Concept		
Category	2, 1151	3.26	0.0389
MCT Status	1, 1151	2.68	0.1021
Ethnic	2, 1151	10.10	0.0001
Category x MCT Status	2, 1143	1.12	0.3272
Category x Ethnic	4, 1143	0.52	0.7228
MCT Status x Ethnic	2, 1143	0.02	0.9800
Category x MCT Status x Ethnic	4, 1139	1.11	0.3501
	Appearance Self-Concept		
Category	2, 1151	8.77	0.0002
MCT Status	1, 1151	1.45	0.2287
Ethnic	2, 1151	20.35	0.0001
Category x MCT Status	2, 1143	0.02	0.9754
Category x Ethnic	4, 1143	1.42	0.2262
MCT Status x Ethnic	2, 1143	0.70	0.4949
Category x MCT Status x Ethnic	4, 1139	0.07	0.5920

	Same Sex Self-Concept		
Category	2, 1151	20.62	0.0001
MCT Status	1, 1151	5.33	0.0211
Ethnic	2, 1151	1.40	0.2471
Category x MCT Status	2, 1143	0.41	0.6647
Category x Ethnic	4, 1143	1.92	0.1054
MCT Status x Ethnic	2, 1143	0.63	0.5342
Category x MCT Status x Ethnic	4, 1139	0.94	0.4426
	Opposite Sex Self-Concept		
Category	2, 1151	4.44	0.0120
MCT Status	1, 1151	1.66	0.1975
Ethnic	2, 1151	11.72	0.0001
Category x MCT Status	2, 1143	0.24	0.7897
Category x Ethnic	4, 1143	1.83	0.1198
MCT Status x Ethnic	2, 1143	0.91	0.4022
Category x MCT Status x Ethnic	4, 1139	0.74	0.5673

Summary of non-academic self-concept. Findings that our three sample groups (RC, EM, and LH) differed on non-academic self-concept domains raise several interesting possibilities. First, self-concept is believed to become more differentiated with age. Our subjects might have been expected to separate their degree of academic success from their non-academic life experiences. That is, academic difficulites would be expected to yield lower academic self-concept scores, but should not exert an effect on the non-academic domains if, in fact, the anticipated differentiation occurs. However, it is possible that academic difficulties did not lead to

depreciations in non-acadmic self-concept domains; rather, the factors that give rise to academic problems also lead to lower rates of success in non-academic self-concept domains as well. The significant multivariate test results confirmed that students differing in category and MCT status differed on non-academic self-concept scores. The design employed in this investigation precludes isolating whether this is due to negative academic experiences generalizing to the non-academic domains or whether EM an LH students and/or those who failed MCTs experience comparable "failure" in non-academic domains, as well.

Attitude Toward School

The MANOVA on dimensions of school attitude scores used four school attitude scales as dependent variables: Mathematics, Social Studies, Science, and Reading/Language attitudes. This analysis is displayed in Table 4.12.

Table 4.12

Results of MANOVA on School Attitude Scales

	$df(df_n, df_d)$	F	p
Category	8, 2298	4.4785	0.0001
MCT Status	4, 1149	1.1582	0.3277
Ethnic	8, 2298	4.3234	0.0001
Category x MCT Status	8, 2282	1.0119	0.4246
Category x Ethnic	16, 3486	1.1052	0.3434
MCT Status x Ethnic	8, 2280	0.5246	0.8388
Category x MCT Status x Ethnic	16, 3474	1.944	0.0133

Category main effect. The multivariate test of differences between the

three categories of students on these variables was statistically significant, $F(8,2298) = 4.48$, $p < .0001$. Univariate tests were used to further examine the significant multivariate test, and the results are presented in Table 4.14. On three of the attitude scales, Reading, Science and Social Studies, the category effect was significant (all $ps < .05$). The univariate test of differences on Math was not significant.

Significant differences ($p < .05$) existed between RC and EM on attitudes toward Reading and Social Studies, where RC students scored significantly higher than did the EM and LH students in Reading and Social Studies (see Table 4.13 for means and standard deviations). In addition, RC and LH groups differed on attitudes toward Reading. In all of these cases, the RC group expressed more favorable attitudes than did the other group. Finally, a significant difference was found between EM and LH subjects on attitudes toward Social Studies, with the LH group having more favorable attitudes. Other contrasts failed to reach significance ($p > .05$).

MCT Status main effect. Students passing or failing the MCT did not differ significantly on their Attitude Toward School score ($p > .10$). As such, it was not necessary to further explore the results of the univariate analyses.

Ethnic main effect. Multivariate test of differences between the three ethnic groups on four dimensions of school attitudes was statistically significant, $F(8,2298) = 4.32$, $p < .0001$. Group means are presented in Table 4.13.

Somewhat unexpectedly, Anglo students had the lowest mean scores on each of the four school attitude scales, while Black and Chicano students had more favorable, and approximately equal, mean scores on all four scales. These

trends were supported by Dunn's test results, which showed that Anglo students differed significantly from both Black and Chicano subjects in attitudes toward Reading ($ps < .05$), with Black and Chicano students expressing more favorable attitudes toward Reading. Anglo subjects also differed significantly ($p < .05$) from Black, but not Chicano, subjects on attitudes toward Math; and they differed significantly from Chicano and Black subjects ($p < .05$) on attitudes toward Social Studies. In all cases, the Anglo subjects expressed less favorable attitudes. All other contrasts failed to reach significance ($p > .05$).

Table 4.13

Means and Standard Deviations of Attitude Toward School Variables, by MCT Status, Category and Ethnicity

	Regular Class				Educationally Marginal				Learning Handicapped			
	Passed		Failed		Passed		Failed		Passed		Failed	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
						Reading						
A'	2.12	.38	2.18	.34	2.07	.32	1.95	.36	2.12	.39	1.98	.36
B'	2.21	.31	2.24	.42	1.94	.55	2.26	.32	2.36	.39	2.16	.31
C'	2.20	.33	2.12	.40	2.10	.35	2.09	.40	1.91	.29	2.17	.37
						Math						
A'	2.04	.44	2.05	.41	2.01	.41	1.87	.42	2.08	.42	1.97	.44
B'	2.05	.41	2.15	.51	1.99	.53	2.19	.43	2.07	.69	2.16	.47
C'	2.03	.42	2.01	.51	2.11	.40	2.01	.45	2.10	.37	2.13	.40
						Science						
A'	2.30	.33	2.32	.36	2.27	.34	2.11	.40	2.30	.34	2.20	.39
B'	2.30	.37	2.32	.35	2.14	.49	2.33	.35	2.47	.50	2.31	.36
C'	2.36	.28	2.30	.35	2.25	.31	2.27	.35	2.44	.34	2.35	.30
						Social Studies						
A'	2.25	.34	2.28	.34	2.25	.32	2.17	.39	2.35	.39	2.23	.37
B'	2.21	.42	2.23	.35	2.07	.32	2.27	.37	2.71	.28	2.35	.36
C'	2.31	.37	2.26	.37	2.17	.38	2.25	.39	2.40	.32	2.36	.34

Note. A' = Anglo; B' = Black; C' = Chicano.

Table 4.14

Results of Univariate Analyses on Survey of School Attitude Scales

	df	F	p
	Attitude Toward Reading		
Category	2, 1152	9.76	0.0001
MCT Status	1, 1152	0.12	0.7282
Ethnic	2, 1152	11.23	0.0001
Category x MCT Status	2, 1144	0.54	0.5837
Category x Ethnic	4, 1144	0.95	0.4335
MCT Status x Ethnic	2, 1144	0.63	0.5309
Category x MCT Status x Ethnic	4, 1140	4.56	0.0012
	Attitude Toward Math		
Category	2, 1152	2.14	0.1183
MCT Status	1, 1152	0.76	0.3845
Ethnic	2, 1152	6.39	0.0017
Category x MCT Status	2, 1144	1.46	0.2330
Category x Ethnic	4, 1144	1.84	0.1196
MCT Status x Ethnic	2, 1144	1.64	0.1938
Category x MCT Status x Ethnic	4, 1140	0.58	0.6783

	Attitude Toward Science		
Category	2, 1152	7.90	0.0004
MCT Status	1, 1152	2.59	0.1079
Ethnic	2, 1152	6.95	0.0010
Category x MCT Status	2, 1144	1.29	0.2754
Category x Ethnic	4, 1143	1.19	0.3156
MCT Status x Ethnic	2, 1143	0.50	0.6064
Category x MCT Status x Ethnic	4, 1140	2.14	0.0737
	Attitude Toward Social Studies		
Category	2, 1152	5.25	0.0054
MCT Status	1, 1152	0.06	0.8056
Ethnic	2, 1152	1.60	0.2017
Category x MCT Status	2, 1144	1.76	0.1721
Category x Ethnic	4, 1144	1.46	0.2137
MCT Status x Ethnic	2, 1144	0.07	0.9292
Category x MCT Status x Ethnic	4, 1140	1.99	0.0942

Summary of Attitude Toward School. One of the major concerns of special educators has been that we provide educational environments and experiences that will enhance the attitudes of mildly handicapped students toward themselves and school. Findings presented here on attitudes toward school are an initial step in attepting to understand the consequences of placements (e.g., special education vs. regular education) on affective outcomes. There is a paucity of empirical evidence on attitudes toward school (MacMillan, et al., 1991).

Attitudes toward school differed for our three sample groups - RC, EM,

and LH. As one might expect, the RC sample characterized by the most successful school careers tended to expresss significantly more favorable attitudes toward Reading and Social Studies than did EM or LH subjects. Univariate tests failed to reveal EM-LH differences with one exception -- on attitudes toward Social Studies. In that instance, the attitudes of the LH group were more favorable. Consider that the EM sample is significantly superior to the LH group in academic competence and they have avoided what some have argued is a stigmatizing experience of being labeled and singled out as different and/or segregated in a special day class. Despite avoiding these experiences and having superior achievement, we were unable to show more favorable attitudes toward three subject matter areas (Reading, Math, and Science) and, in fact found EM subjects to express significantly less favorable attitudes toward Social Studies. If mainstreaming is expected to result in more favorable attitudes toward school, this pattern of results on nonhandicapped students in the lowest quartile of achievement gives reason for concern.

Results on Ethnic group differences in attitudes toward school were also somewhat surprising. Anglo subjects across all three category groups expressed significantly less favorable attitudes than at least one of the ethnic minority groups on three of the four scales (Science being the exception). These attitudes are not congruent with measured achievement of the groups (presented in Tables 4.17 through 4.19 later in this chapter). The explanation for the poorer attitudes expressed by Anglo subjects, despite comparable levels of acheivement, warrants further investigation. MacMillan et al, (1992) suggested that level of performance academically is compared to certain expectations. That is, Anglo students know that, in general, Black

students do not achieve as well as Anglo students, and in our samples Black and White students achieved at the same level. It is conceivable that Anglo students perceived the situation (i.e., not exceeding the level of Black students) as indicative of "poor" performance and expressed attitudes toward subject matter areas in keeping with the personalized evaluation. Similarly, Black students may perceive the fact that they are doing "as well as" Anglo students as indicative of "good" performance and express attitudes accordingly.

MCT Performance, Student Characteristics, & Behavior

This section will present descriptive data generated in response to the research questions posed in Aim II of the initial proposal addressing the proportions of students passing the MCT, sorted by ethnicity, and category. Specifically, we hypothesized that students repeatedly failing the MCT will more frequently transfer to a neighboring school district than those who have passed the MCT. Such transfers will be more frequent from districts with "harder" MCTs than from districts with "easier" MCTs. (The level of difficulty of MCTs was discussed in the previous chapter).

We also hypothesized that MCT failure would result in high rates of school dropout among EM subjects. Were that to happen, one might find very low rates of denying diplomas at the time of graduation due to MCT failure, because the very students who failed the MCT (a basis for denying a diploma) had dropped out before completing academic coursework and being a candidate for graduation. Both transfer and dropout behaviors were expected to be higher among EM students than LH or RC students. On-time graduation (age was used as a proxy for the first cohort of students) was used in determining differences between students passing MCTs and those failing MCTs.

Four-way frequency analyses were performed to develop hierarchical log-linear models of the effects of MCT status, ethnic group, and category on dropout, transfer, and over-age status. Stepwise selection by simple deletion of effects using BMDP4F produced a model that included all first-order effects as well as all possible two-, three-, and four-way marginal associations.

Dropout

The two-way interaction of Category and Dropout had a significant likelihood ratio $\chi^2(2,N = 1188) = 17.52$, $p < 0.0001$, with 11 percent of the RCs dropping out, 15 percent of the EMs and 20 percent of the LHs. The two-way association of MCT status and Dropout was significant $\chi^2(4,N = 1188) = 97.14$, $p < 0.0001$ indicates that passage of the MCT significantly influenced the rate of dropping out of school. As hypothesized, there was a higher dropout rate for students failing the MCT than for students passing the MCT. Specifically, 38 percent (275/723) of the students failing the MCT dropped out, whereas only thirteen percent (59/465) of students passing the MCT dropped out. The likelihood χ^2 ratio tests for ethnicity and dropout did not indicate significant effects involving these variables.

Table 4.15

Frequencies by Ethnic, Category, MCT Status and Dropout

			Anglo	Black	Chicano	Total
Failed	Persist	RC	64	40	43	147
		EM	88	22	57	167
		LH	88	12	34	134
	Total		240	74	134	448
	Dropout	RC	45	14	26	85
		EM	51	15	38	104
		LH	47	15	24	86
	Total		143	44	88	275
Passed	Persist	RC	143	60	77	280
		EM	55	8	34	97
		LH	20	3	6	29
	Total		218	71	117	406
	Dropout	RC	17	12	12	41
		EM	4	0	6	10
		LH	4	1	3	8
	Total		25	13	21	59

Note. Total of the observed frequency table is 1188. All cases had complete data for this table.

Table 4.15a

Frequencies of the two-way significant interaction of Category and Dropout

	RC	EM	LH	Total
Persist	427	264	163	854
Dropout	126	114	94	334

Table 4.15b

Frequencies of the two-way significant interaction of MCT Status and Dropout

	Pass	Fail	Total
Persist	406	448	854
Drop	59	275	334

Table 4.16

Chi-Square Results of Analyses on Ethnic, Category, MCT Status and Dropout

Effect	D.F.	Chi-Square	p-value	Iteration
Ethnic (ETH)	2	232.76	0.0000	-
Category (CAT)	2	111.95	0.0000	-
Dropout (DROP)	1	235.50	0.0000	-
MCT Status (MCT)	1	56.48	0.0000	-
ETH x CAT	4	31.65	0.0000	2
ETH x DROP	2	1.33	0.5136	2
ETH x MCT	2	0.63	0.7302	2
CAT x DROP	2	17.52	0.0002	2
CAT x MCT	2	175.94	0.0000	2
DROP x MCT	1	97.14	0.0000	2

Note. Association option selected for all terms of order less than or equal to four. No significant three- or four-way interactions.

Transfer

A four-way frequency analysis was performed to develop a hierarchical log-linear model of transfer, MCT status, ethnic group, and category. Results indicated that students failing the MCT were more likely to transfer to another school, $\chi^2(4,N = 1188) = 57.20$, $p < 0.0001$ (see Table 4.18). There was no difference in the transfer behavior among categories; however, there was a difference among the ethnic groups' transfer behavior $\chi^2(4,N = 1188) = 10.86$, $p < 0.0015$. Thirty percent (194/626) of the Anglo group transfered, 32 percent (66/202) of the Black students and only 22 percent (80/360) of the Chicano students transfered.

Table 4.17

Frequencies by Ethnic, Category, MCT Status and Transfer

			Anglo	Black	Chicano	Total
No Transfer	Failed	RC	70	27	48	145
		EM	80	19	70	169
		LH	83	20	43	146
	Total		233	66	161	460
	Passed	RC	132	64	79	275
		EM	48	4	32	84
		LH	19	2	8	29
	Total		199	70	119	388
Transfer	Failed	RC	39	27	21	87
		EM	59	18	25	102
		LH	52	7	15	74
	Total		150	52	61	263
	Passed	RC	28	8	10	46
		EM	11	4	8	23
		LH	5	2	1	8
	Total		44	14	19	77

Note. Total of the observed frequency table is 1188. All cases had complete data for this table.

Table 4.17a

Frequencies of the two-way significant interaction of MCT Status and Transfer

	Anglo	Black	Chicano	Total
No Transfer	432	136	280	848
Transfer	194	66	80	340

Table 4.17b

Frequencies of the two-way significant interaction of MCT Status and Transfer

	Pass	Fail	Total
No Transfer	388	460	848
Transfer	77	263	340

Table 4.18

Chi-Square Results of Analyses on Ethnic, Category, MCT Status and Transfer

Effect	D.F.	Chi-Square	p-value	Iteration
Ethnic (ETH)	2	232.76	0.0000	-
Category (CAT)	2	111.95	0.0000	-
MCT Status (MCT)	1	56.48	0.0000	-
Transfer (TRAN)	1	224.38	0.0000	-
ETH x CAT	4	31.65	0.0000	2
ETH x MCT	2	0.63	0.7301	2
ETH x TRAN	2	10.86	0.0044	2
CAT x MCT	2	175.94	0.0000	2
CAT x TRAN	2	10.74	0.0047	2
MCT x TRAN	1	57.20	0.0000	2

Note. Association option selected for all terms of order less than or equal to four. No significant three- or four-way interactions.

Over-Age

As a proxy for on-time graduation, a variable called over-age was created. A student was considered to be over the expected age for graduation if s/he was 18 before November of what should be his/her senior year. This definition was derived from the fact that most districts use November as the cutoff point for Kindergarten registration. If the child is not five by November, the student must postpone enrollment until the following year. A nineteen year-old in the senior class is clearly older than his classmates, and probably experienced at least one grade retention during his academic career. (The reader should be reminded that these analyses were performed using only the first two years of data. As such, "over-age" is based on a student's age at the end of what should be his freshman year in high school.)

The two-way association between over-age and ethnicity yielded a significant effect, with a likelihood ratio $\chi^2(4,N = 1188) = 14.28$, $p < 0.0015$ (see Table 4.20). This effect indicated that there were ethnic group differences in the over-age status. These differences are displayed in Table 4.19, a three-way contingency table. Specifically, Chicano students exhibited a higher rate of over-age status (27 percent, or 96 of 360), whereas about 15 percent of Black students (31 of 202) and 16 percent of Anglo students (101 of 616) were over-age.

Differences in the age distribution were also based on category, $\chi^2(4,N = 1188) = 98.16$, $p < 0.0000$), with the largest percent of over-age students in the LH (34 percent, 88/257) and EM (27 percent, 104/378) categories being over-age. In striking contrast, only 8 percent (46/553) of the RCs were over-age.

The significant likelihood ratio between over-age and MCT status, $\chi^2(4,N$

= 1188) = 46.13, $p < 0.0000$, indicates an age difference in those passing versus failing the MCT. The obvious direction of this difference is the over-age student failed the MCT more (79 percent, 189/238) than the at-grade aged student (56 percent, 534/950).

Table 4.19

Frequencies by Ethnic, Category, Over-Age and MCT Status

			Anglo	Black	Chicano	Total
On Time	Failed	RC	98	50	58	206
		EM	104	26	57	187
		LH	94	14	33	141
	Total		296	90	148	534
	Passed	RC	153	70	78	301
		EM	48	8	31	87
		LH	18	3	7	28
	Total		219	81	116	416
Over-Age	Failed	RC	11	4	11	26
		EM	35	11	38	84
		LH	41	13	25	79
	Total		87	28	74	189
	Passed	RC	7	2	11	20
		EM	11	0	9	20
		LH	6	1	2	9
	Total		24	3	22	49

Note. Total of the observed frequency table is 1188. All cases had complete data for this table.

Table 4.19a

Frequencies of the two-way significant interaction of Over-Age and Category

	RC	EM	LH	Total
On-Time	507	274	169	950
Over-Age	46	104	88	238

Table 4.19b

Frequencies of the two-way significant interaction of Over-Age and MCT-Status

	Pass	Fail	Total
On-Time	416	534	950
Over-Age	49	189	238

Table 4.19c

Frequencies of the two-way significant interaction of Over-Age and Ethnicity

	Anglo	Black	Chicano	Total
On-Time	515	171	264	950
Over-Age	111	31	96	238

Table 4.20

Chi-Square Results of Analyses on Ethnic, Category, MCT Status and Over-Age

Effect	D.F.	Chi-Square	p-value	Iteration
Ethnic (ETH)	2	2332.76	0.0000	-
Category (CAT)	2	111.95	0.0000	-
MCT Status (MCT)	1	56.48	0.0000	-
Over-Age (AGE)	1	456.85	0.0000	-
ETH x CAT	4	31.65	0.0000	2
ETH x MCT	2	0.63	0.7302	2
ETH x AGE	2	14.28	0.0008	2
CAT x MCT	2	175.94	0.0000	2
CAT x AGE	2	98.16	0.0000	2
MCT x AGE	1	46.13	0.0000	2

Note. Association option selected for all terms of order less than or equal to four. No significant three- or four-way interactions.

Academic Achievement

A third aim of the project was to compare ethnicity and category on measured standardized achievement and MCT performance.

Achievement on Standardized Tests

A significant ethncity X category interaction was predicted, with minority students in LH and EM samples expected to score significantly below majority students in these two samples, and with no differences in measured achievement between ethnic groups in the RC sample. In order to test these hypothesized group differences, orthogonal contrasts of interest were specified prior to the analyses. The first of these contrasts the RC group with the LH and EM groups combined. The second a priori contrast looks at the LH and EM groups, holding the RCs constant. The first MANOVA was based on

Reading and Math achievement test scores (see Table 4.22).

Category Main Effect. The multivariate test of differences between the three categories of students was statistically significant, $F(4,1196) = 108.02$, $p < .0001$. This overall test was followed up with multivariate tests of the two orthogonal contrasts of interest. The first contrast, comparing the RC students to the two groups of low achieving students was significant, $F(2,598) = 254.95$, $p < .0001$. The second contrast, comparing the EM and LH students, was also significant, $F(2,598) = 8.81$, $p < .0002$.

Tables 4.23 and 4.24 present univariate results revealing the difference in reading and math achievement scores between RC and the two low achieving groups to be significant, $F(1,599) = 243.72$, $p < .0001$, and the difference in scores between the EM and LH groups to also be significant, $F(1,599) = 9.97$, $p < .002$. The mean reading achievement test scores fell in an easily interpretable trend, as the RC, EM, and LH groups had mean NCE scores of 53.30, 31.05, and 23.55, respectively, on the reading achievement test (see Table 4.21).

A similar pattern of results was obtained on mathematics achievement test scores. The contrast between the RC and the combined EM and LH groups was significant, $F(1,599) = 463.98$, $p < .0001$, as was the contrast between the EM and LH groups, $F(1,599) = 15.03$, $p < .0001$. As with reading achievement, the RC group ($M = 64.68$) showed the highest level of math achievement, the EM group ($M = 38.44$) a moderate level, and the LH group ($M = 30.44$) the lowest level of math achievement.

Ethnic Main Effect. The multivariate test of differences between the three ethnic groups in academic achievement was of borderline significance, $F(4,1196) = 2.35$, $p < .06$. The mean scores on the achievement tests, found in

Table 4.21, revealed Anglo and Black students had similar mean levels of achievement in reading (Anglo $\underline{M}$ = 41.40; Black $\underline{M}$ = 41.83) and math (Anglo $\underline{M}$ = 49.88; Black $\underline{M}$ = 49.16), although both groups scored higher than did Chicano students (reading $\underline{M}$ = 38.06; math $\underline{M}$ = 47.77). These differences were confirmed via multivariate test, with a nonsignificant difference between Anglo and Black students, $\underline{F}(2,598) = 0.36$, $\underline{p} > .65$, but a significant difference between the combined Anglo/Black group and Chicano students, $\underline{F}(2,598) = 4.18$, $\underline{p} < .02$. Univariate results on both reading and math achievement test scores also confirmed these differences, with nonsignificant differences between Anglo and Black students (both $\underline{p}$s > .50), but significant differences between the combined Anglo/Black and Chicano students (both $\underline{p}$s < .05).

Table 4.21

Mean Scores on Achievement Tests

	RC	EM	LH
Reading	52.95	31.15	24.14
Math	63.72	38.46	31.51

Table 4.22

Results of MANOVA on Achievement Test

	$df(df_n, df_d)$	F	p
Category	4, 1206	108.49	0.0000
RC vs EM & LH (D1)	2, 603	255.84	0.0000
EM vs LH (D2)	2, 603	9.30	0.0001
Gender	2, 603	0.51	0.6010
Ethnic	2, 1206	2.23	0.0636
Anglo vs Black & Chicano (MAJ)	2, 603	3.37	0.0351
Black vs Chicano (MIN)	2, 603	0.54	0.5848
Category x Gender	4, 1190	0.66	0.6184
D1 x Gender	2, 595	1.28	0.2782
D2 x Gender	2, 595	0.14	0.8667
Category x Ethnic	8, 1190	0.46	0.8487
D1 x Ethnic	4, 1190	0.22	0.9278
D2 x Ethnic	4, 1190	0.69	0.5986
Category x MAJ	4, 1190	0.51	0.7257
Category x MIN	4, 1190	0.34	0.8513
D1 x MAJ	2, 595	0.36	0.6957
D2 x MAJ	2, 595	0.66	0.5147
D1 x MIN	2, 595	0.10	0.9094
D2 x MIN	2, 595	0.58	0.5577
Gender x Ethnic	4, 1190	1.92	0.1047
Gender x MAJ	2, 595	3.17	0.0427
Gender x MIN	2, 595	0.45	0.6399
Category x Ethnic x Gender	8, 1182	0.23	0.9234
D1 x Ethnic x Gender	4, 1182	0.23	0.9234
D2 x Ethnic x Gender	4, 1182	1.53	0.1907
D1 x MAJ x Gender	2, 591	0.28	0.7537
D2 x MAJ x Gender	2, 591	0.32	0.7236
D1 x MIN x Gender	2, 591	0.13	0.8747

D2 x MIN x Gender	2, 591	2.78	0.0627
Category x Gender x MAJ	4, 1182	0.30	0.8759
Category x Gender x MIN	4, 1182	1.46	0.2136

Note: MAJ= Anglo versus Blacks and Chicanos; MIN = Blacks versus Chicanos, holding Anglo's constant.

Table 4.23

Results of Univariate Analyses on Reading Achievement Test

	$df(df_n, df_d)$	F	p
Category	2, 595	123.26	0.0000
RC vs EM & LH (D1)	2, 603	246.35	0.0000
EM vs LH (D2)	2, 603	10.76	0.0011
Gender x Ethnic	4, 1190	1.92	0.1047
Gender x MAJ	2, 595	3.17	0.0427
Category x Ethnic x Gender	8, 1182	0.23	0.9234
D2 x Ethnic x Gender	4, 1182	1.53	0.1907

Table 4.24

Results of Univariate Analyses on Math Achievement Test

	$df(df_n, df_d)$	F	p
Category	2, 5	232.37	0.0000
RC vs EM & LH (D1)	1, 5	464.73	0.0000
EM vs LH (D2)	1, 5	15.69	0.0001
Ethnic	2, 5	3.40	0.0341
Anglo vs Black & Chicano (MAJ)	1, 5	5.98	0.0147

Summary of Achievement Tests. As would be expected, category is a good predictor of achievement test performance. Both Math and Reading achievement differs significantly by category, with the scores of the groups falling in a predictable pattern with the RCs having the highest scores, EMs next highest, and LHs lowest scores. The orthogonal contrasts yielded support for both a priori contrasts for Reading and Math achievement. RCs differed from EMs and LHs, and EMs differed from LHs on both achievement tests.

Minimum Competency Test Scores

Category Main Effect. The multivariate test of differences between the three categories of students was statistically significant, $F(4,608) = 14.15$, $p < .0000$. This overall test was followed up with multivariate tests of the two orthogonal contrasts of interest. The first contrast, comparing the RC students to the two groups of low achieving students was significant, $F(2,304) = 28.17$, $p < .0000$. The second contrast, comparing the EM and LH students, was also significant, $F(2,304) = 5.11$, $p < .01$ (see Table 4.26).

The univariate results reveal the difference in Reading MCT scores between RC and the combined EM and LH groups to be significant, $F(1,5) = 53.37$, $p < .0001$, as was the comparison between the EM and LH groups on

Reading MCT scores, $F(1,5) = 8.03$, $p < .01$ (see Tables 4.27 and 4.28). The mean Reading MCT scores fell in an easily interpretable trend, as the RC, EM, and LH groups had mean NCE scores of 57.54, 43.94, and 34.85, respectively, on the Reading MCT (see Table 4.25).

A similar pattern of results was obtained on Mathematics MCT scores. The contrast between the RC and the combined EM and LH groups was significant, $F(1,5) = 40.27$, $p < .0001$, as was the contrast between the EM and LH groups, $F(1,5) = 9.22$, $p < .01$. As with Reading MCT, the RC group ($M = 51.08$) showed the highest Math MCT scores, the EM group ($M = 38.92$) a moderate level, and the LH group ($M = 28.29$) the lowest Math MCT scores (see Table 4.25).

Ethnic Main Effect. The multivariate test of differences between the three ethnic groups in MCT was not significant, $F(4,608) = 0.41$, $p > .10$ (see Table 4.26).

Interactions. As category was the only main effect on MCT that was significance, it follows that no interactions between the main effects was significant (see Table 4.26).

Table 4.25

Mean Scores on Minimum Competency Tests

	RC	EM	LH
Reading	57.54	43.94	34.85
Math	51.08	38.92	28.29

Table 4.26

Results of MANOVA on Minimum Competency Test Scores

	$df(df_n, df_d)$	F	p
Category	4, 608	14.15	0.0000
RC vs EM & LH (D1)	2, 304	28.17	0.0000
EM vs LH (D2)	2, 304	5.11	0.0001
Gender	2, 304	0.78	0.6010
Ethnic	4, 608	0.41	0.0636
Anglo vs Black & Chicano (MAJ)	2, 603	3.37	0.0351
Black vs Chicano (MIN)	2, 603	0.54	0.5848
Category x Gender	4, 592	0.82	0.6184
D1 x Gender	2, 296	0.25	0.2782
D2 x Gender	2, 296	1.44	0.8667
Category x Ethnic	8, 592	1.53	0.8487
D1 x Ethnic	4, 592	2.13	0.9278
D2 x Ethnic	4, 592	0.87	0.5986
Category x MAJ	4, 592	0.99	0.7257
Category x MIN	4, 592	2.05	0.8513
D1 x MAJ	2, 296	0.18	0.6957
D2 x MAJ	2, 296	1.79	0.5147
D1 x MIN	2, 296	4.04	0.9094
D2 x MIN	2, 296	0.09	0.5577
Gender x Ethnic	4, 592	0.38	0.1047
Gender x MAJ	2, 296	0.71	0.0427
Gender x MIN	2, 296	0.10	0.6399

Category x Ethnic x Gender	8, 584	1.52	0.9234
D1 x Ethnic x Gender	4, 584	1.09	0.9234
D2 x Ethnic x Gender	4, 584	0.96	0.1907
D1 x MAJ x Gender	2, 292	1.67	0.7537
D2 x MAJ x Gender	2, 292	0.48	0.7236
D1 x MIN x Gender	2, 292	0.47	0.8747
D2 x MIN x Gender	2, 292	1.49	0.0627
Category x Gender x MAJ	4, 584	1.07	0.8759
Category x Gender x MIN	4, 584	0.98	0.2136

Note: MAJ= Anglo versus Blacks and Chicanos; MIN = Blacks versus Chicanos, holding Anglo's constant.

Table 4.27

Results of Univariate Analyses on Reading Minimum Competency Test

	$df(df_n, df_d)$	F	p
Category	2, 5	27.44	0.0001
RC vs EM & LH (D1)	1, 5	53.37	0.0001
EM vs LH (D2)	1, 5	8.03	0.0049
Category x Ethnic			
D1 x Ethnic	2, 13	4.12	0.0172
Category x MIN	2, 13	4.08	0.0179
D1 x MIN	2, 13	8.02	0.0050

Note. Only those effects significant at the .05 level or greater are reported.

negative dimensions of school life.

Performance on the MCT greatly impacts the types of courses selected. Students passing the MCT enroll in more elective courses and fewer remedial courses. This is not surprising, as failure of the MCT usually relegates a student to remedial courses until he passes the MCT. While not surprising, this result is disturbing in light of the fact that it is probably the very students not passing the MCT that will enter the job market out of high school. Not only are these students most likely to have failed repeated remedial courses (Balow & Schwager, 1990), they are the students who could most benefit from elective vocational courses.

Because of the differing ability levels of the MCT between districts, it was hypothesized that students failing the MCT in one district may transfer to a district that has an easier test. Indeed, those students failing the MCT did transfer up to ten times (LH group) more frequently than those passing the MCT. However, it is not clear if it is those students who transfer in and out of districts who fail the MCT, or if students transfer out of a district because of MCT failure. Related to this concern is the question of whether those students (or their parents) failing the MCT would be savvy to the different difficulty levels of the MCTs in neighboring districts.

A significant relationship between dropping out of high school and MCT performance was found. Depending on the academic grouping of a student, the dropout behavior of students failing the MCT was up to ten times (LH group) more frequent than for those students passing the MCT. Again, the same caveat applies to this group of students as to those transferring out of districts.

In terms of school affect, students failing the MCT are absent more often, and dropout more frequently. Students failing the MCT have lower

General and Academic self-concept as well as lowered self-concept in math, honesty and same-sex relations. While failing the MCT does not appear to affect students' attitude toward school as measured by the Survey of School Attitudes, the fact that those failing the MCT do not come to school as often and dropout of school more frequently could be considered as a proxy for an overall, negative school attitude.

Certainly results of these analyses provide only the beginning to many related inquiries into the school life of students of all academic ability levels, and the consequences of resulting school failure.

References

Balow, I. H., & Schwager, M. (1990, February). Retention in grade: A failed procedure. A Report presented to members of the California Educational Research Cooperative. University of California, Riverside.

Humphreys, L. G., & Fleishman, A. (1974). Pseudo-orthogonal and other analysis of variance designs involving individual-differences variables. Journal of Educational Psychology, 66, 464-472.

MacMillan, D. L., Widaman, K. F., Balow, I. H., Hemsley, R. E., and Little, T. D. (1992). Differences in adolescent school attitudes as a function of academic level, ethnicity, and gender. Learning Disability Quarterly, 15, 39-50.

Widaman, K. F., MacMillan, D. L., Hemsley, R. E., Little, T. D., and Balow, I. H. (1992). Differences in adolescents' self-concept as a function of academic level, ethnicity and gender. American Journal on Mental Retardation. 96, (4). 387-405.

Sample Data Collection Sheet

(1) ____ ____ ____:Project ID#

(2) ____ ____ ____:First Initial of Subject's First, Middle, & Last Name

(3) ____:Cohort

(4) ____ ____ / ____ ____ / ____ ____:Birthdate (Month / Day / Year)

(5) ____ ____ - ____ ____:Present School and District (Indicate Attrition Status Here, If Appropriate)

(6) ____ ____:Grade

(7) ____:Repeat / Retained This Grade

(8) ____:Placement (Indicates any Special Services Received This Academic Year)

(9) ____ ____ ____ - ____ ____ - ____ ____:Attendance (Total Days Absent For Year 1st Sem.-2nd Sem.)

(10) ____:Achievement Test

(11) ____:District Where Test Was Taken

(12) ____ ____ - ____:Achievement Test Level and Battery

(13) ____ ____:Publication Date of Achievement Test

(14) ____:Type of Score Recorded for Achievement Test

(15) ____ ____ ____:Math Score on Achievement Test

(16) ____ ____ ____:Reading Score on Achievement Test

(17) ____ ____:Competency Test

(18) ____:District Where Competency Test Was Taken

(19) ____ ____ - ____:Competency Test Level and Battery (20) ____:Type of Score Recorded for Competency Test

(21) ____ ____ ____:Cutoff for Math Score on Competency Test

(22) ____ ____ ____ - ____:Math Score on Competency Test - Indicate if Subject Passed or Failed MCT

(23) ____ ____ ____:Cutoff for Reading Score on Competency Test

(24) ____ ____ ____ - ____:Reading Score on Competency Test - Indicate if Subject Passed or Failed MCT

(25) ____ ____ - ____:Math Grades (1st semester - 2nd semester - Indicate if Remedial Class. If Not, Score ----)

(26) ____ ____:Math Citizenship Grades (1st semester - 2nd semester)

(27) ____ ____ - ____:Science Grades (1st semester - 2nd semester - Indicate if Remedial Class. If Not, Score ----)

(28) ____ ____:Science Citizenship Grades (1st semester - 2nd semester)

Datasheet 12 (1985-1986) page 2

(29) ____ ____-____:Soc. Stud./History Grades (1st semester - 2nd semester - Indicate if Remedial Class. If Not, Score)

(30) ____ ____:Soc. Stud./History Citizenship Grades (1st semester - 2nd semester)

(31) ____ ____:Gen. Citizenship (Overall citizenship for 1st semester - 2nd semester)

(32) ____:Teacher Comments on Mental Ability

(33) ____:Teacher Comments on General Academic Competence

(34) ____:Teacher Comments on Academic Competence in Math

(35) ____:Teacher Comments on Academic Competence in Reading

(36) ____:Teacher Comments on Academic Competence in English

(37) ____:Teacher Comments on Student's Class Work Habits

(38) ____:Teacher Comments on Class Adjustment/Affective behavior

(39) ____:Teacher Comments About Student's Chronic Absenteeism

(40) ____:Teacher Recomended Placement for Next Year

(41) ____ ____ ____ ____ ____ ____:Diagnostic Problem (Speech, Motor, Medical, Visual, Hearing, Other)

(42) ____ ____ ____ ____ ____:Special Services Received (Speech Therapy, Medication, Motor, Tutoring, Other)

(43) ____ ____ ____ ____:Extracurricular Activities (Academic-Athletic-Interest/Hobby-Social)

(44) ____ ____:Discipline / Suspension Recommendation (Indicate # of Recommendations. If None, Score With ----)

(45) ____ ____:Referral for Anti-Teacher / Anti-Authority Behavior (Indicate # of Referrals. If None, Score With ----)

(46) ____ ____:Referral for Conflict With Other Students (Indicate # of Referrals. If None, Score With ----)

(47) ____ ____:Referral for Maladjustment to Classroom, Not Involving Others (Indicate # of Referrals. If None, Score ---)

(48) ____ ____:Referral to Any School Personnel for Academic Problems (Indicate # of Recommendations. If None, Score ----)

(49) ____ ____:Referral for Substance Abuse (Indicate # of Referrals. If None, Score With ----)

(50) ____ ____:Referral for Behavior Unspecified Above, but Related (Indicate # of Referrals. If None, Score With ----)

(51) ____ ____:Special Achievements, e.g. Awards, Leadership etc. (Indicate # of Referrals. If None, Score With ---)

(52) ____ ____:Parental Request for Child to be Removed From Special Ed. (Indicate # of Referals. If None, Score With ---)

(53) ____ ____:Parental Request for Child to be Returned to Special Ed. (Indicate # of Referals. If None, Score With ---)

(54) ____ ____:Parental Request for Information From School About Child (Indicate # of Referals. If None, Score ----)

(55) _____ ____:Parental Request for School Conference (Indicate # of Referals. If None, Score With ----)

(56) ____ - ____ ____ ____:Type of IQ Test Administered - IQ Score

School _____ Name _____

SELF DESCRIPTION QUESTIONAIRE II

ID # _____

THIS IS A CHANCE TO LOOK AT YOURSELF. IT IS NOT A TEST. THERE ARE NO RIGHT ANSWERS AND EVERYONE WILL HAVE DIFFERENT ANSWERS. BE SURE THAT YOUR ANSWERS SHOW HOW YOU FEEL ABOUT YOURSELF. PLEASE DO NOT TALK ABOUT YOUR ANSWERS WITH ANYONE ELSE. WE WILL KEEP YOUR ANSWERS PRIVATE AND NOT SHOW THEM TO ANYONE. THE PURPOSE OF THIS STUDY IS TO SEE HOW PEOPLE DESCRIBE THEMSELVES.

WHEN YOU ARE READY TO BEGIN, PLEASE READ EACH SENTENCE AND DECIDE YOUR ANSWER. (YOU MAY READ QUIETLY TO YOURSELF IF THEY ARE READ ALOUD TO YOU.) THERE ARE SIX POSSIBLE ANSWERS FOR EACH QUESTION -- "TRUE", "FALSE", AND FOUR ANSWERS IN BETWEEN. THERE ARE SIX BOXES NEXT TO EACH SENTENCE, ONE FOR EACH OF THE ANSWERS. THE ANSWERS ARE WRITTEN AT THE TOP OF THE BOXES. CHOOSE YOUR ANSWER TO A SENTENCE AND PUT A TICK (✓) IN THE BOX UNDER THE ANSWER YOU CHOOSE. DO NOT SAY YOUR ANSWER ALOUD OR TALK ABOUT IT WITH ANYONE ELSE.

BEFORE YOU START THERE ARE THREE EXAMPLES BELOW. I HAVE ALREADY ANSWERED TWO OF THE THREE SENTENCES TO SHOW YOU HOW TO DO IT. IN THE THIRD ONE YOU MUST CHOOSE YOUR OWN ANSWER AND PUT IN YOUR OWN TICK (✓).

	FALSE	MOSTLY FALSE	MORE FALSE THAN TRUE	MORE TRUE THAN FALSE	MOSTLY TRUE	TRUE
1. I LIKE TO READ COMIC BOOKS						✓

(I PUT A TICK IN THE BOX UNDER THE ANSWER "TRUE". THIS MEANS THAT I REALLY LIKE TO READ COMIC BOOKS. IF I DID NOT LIKE TO READ COMIC BOOKS VERY MUCH, I WOULD HAVE ANSWERED "FALSE" OR "MOSTLY FALSE".)

2. IN GENERAL, I AM NEAT & TIDY.			✓			

(I ANSWERED "MORE FALSE THAN TRUE" BECAUSE I AM DEFINITELY NOT VERY NEAT, BUT I AM NOT REALLY MESSY EITHER.)

3. I LIKE TO WATCH T.V.						

(FOR THIS SENTENCE YOU HAVE TO CHOOSE THE ANSWER THAT IS BEST FOR YOU. FIRST YOU MUST DECIDE IF THE SENTENCE IS "TRUE" OR "FALSE" FOR YOU, OR SOMEWHERE IN BETWEEN. IF YOU REALLY LIKE TO WATCH T.V. A LOT YOU WOULD ANSWER "TRUE" BY PUTTING A TICK IN THE LAST BOX. IF YOU HATE WATCHING T.V. YOU WOULD ANSWER "FALSE" BY PUTTING A TICK IN THE FIRST BOX. IF YOU DO NOT LIKE T.V. VERY MUCH, BUT YOU WATCH IT SOMETIMES YOU MIGHT DECIDE TO PUT A TICK IN THE BOX THAT SAYS "MOSTLY FALSE" OR THE BOX FOR "MORE FALSE THAN TRUE".

IF YOU WANT TO CHANGE AN ANSWER YOU HAVE MARKED YOU SHOULD CROSS OUT THE TICK AND PUT A NEW TICK IR ANOTHER BOX ON THE SAME LINE. FOR ALL THE SENTENCES BE SURE THAT YOUR TICK IS ON THE SAME LINE AS THE SENTENCE YOU ARE ANSWERING. YOU SHOULD HAVE ONE ANSWER AND ONLY ONE ANSWER FOR EACH SENTENCE. DO NOT LEAVE OUT ANY SENTENCES, EVEN IF YOU ARE NOT SURE WHICH BOX TO TICK.

IF YOU HAVE ANY QUESTIONS HOLD UP YOUR HAND. OTHERWISE TURN OVER THE PAGE AND BEGIN.

	MOSTLY FALSE	MORE FALSE THAN TRUE	MORE TRUE THAN FALSE	MOSTLY TRUE	TRUE
1. ENGLISH IS ONE OF MY BEST SUBJECTS.					
2. I HATE THINGS LIKE SPORT, GYM, AND DANCE.					
3. BOYS FIND ME BORING.					
4. PEOPLE CAN REALLY COUNT ON ME TO DO WHAT IS RIGHT.					
5. MY PARENTS UNDERSTAND ME.					
6. WHEN I DO A JOB I DO IT WELL.					
7. I LOOK FORWARD TO MATHEMATICS CLASSES.					
8. I FIND IT DIFFICULT TO MEET GIRLS I LIKE.					
9. I AM HAPPY MOST OF THE TIME.					
10. IF I WORK REALLY HARD I COULD BE ONE OF THE BEST STUDENTS IN MY SCHOOL YEAR.					
11. OTHER PEOPLE THINK I AM GOOD LOOKING.					
12. I HAVE A POOR VOCABULARY.					
13. I ENJOY THINGS LIKE SPORTS, GYM & DANCE					
14. I'M UNCOMFORTABLE BEING AFFECTIONATE WITH MEMBERS OF THE OPPOSITE SEX.					
15. I ALWAYS TELL THE TRUTH.					
16. MY PARENTS TREAT ME FAIRLY.					
17. SOMETIMES I THINK THAT I AM NO GOOD AT ALL.					
18. I HATE MATHEMATICS.					
19. GIRLS OFTEN MAKE FUN OF ME.					
20. I USUALLY LOOK ON THE GOOD SIDE OF THINGS.					
21. I AM STUPID IN MOST SCHOOL SUBJECTS.					
22. I HAVE A NICE LOOKING FACE.					
23. WORK IN ENGLISH CLASSES IS EASY FOR ME.					
24. I'M TERRIBLE AT EVERY SPORT I HAVE EVER TRIED.					
25. I AM POPULAR WITH BOYS.					
26. I SOMETIMES TAKE THINGS THAT BELONG TO OTHER PEOPLE.					
27. MY PARENTS REALLY LOVE ME A LOT.					
28. I CAN'T DO ANYTHING RIGHT.					
29. I DO BADLY IN TESTS OF MATHEMATICS.					
30. I AM POPULAR WITH GIRLS.					
31. I AM OFTEN DEPRESSED AND DOWN IN THE DUMPS.					
32. MOST SCHOOL SUBJECTS ARE JUST TOO HARD FOR ME.					
33. I AM GOOD LOOKING.					
34. I LOOK FORWARD TO ENGLISH CLASSES.					
35. I TRY TO GET OUT OF SPORTS & PHYSICAL EDUCATION CLASSES WHENEVER I CAN.					
36. MOST BOYS WANT ME TO BE THEIR FRIEND.					
37. I OFTEN TELL LIES.					
38. MY PARENTS PUNISH ME MORE SEVERELY THAN I DESERVE.					
39. I HATE MYSELF.					
40. I OFTEN NEED HELP IN MATHEMATICS.					
41. MOST GIRLS TRY TO AVOID ME.					
42. I AM A CALM PERSON.					
43. I LEARN THINGS QUICKLY IN MOST SCHOOL SUBJECTS.					
44. THERE ARE A LOT OF THINGS ABOUT THE WAY I LOOK THAT I WOULD LIKE TO CHANGE.					
45. I GET GOOD MARKS IN ENGLISH.					
46. I AM A SLOW RUNNER.					
47. I FIND IT DIFFICULT TO MEET BOYS I LIKE.					
48. HONESTY IS VERY IMPORTANT TO ME.					
49. IF I HAVE CHILDREN OF MY OWN, I WANT TO BRING THEM UP LIKE MY PARENTS RAISED ME.					
50. OVERALL, I AM NO GOOD.					
51. MATHEMATICS IS ONE OF MY BEST SUBJECTS.					
52. PEOPLE OF THE OPPOSITE SEX THAT I LIKE DON'T LIKE ME.					
53. I OFTEN FEEL CONFUSED AND MIXED UP.					
54. I ENJOY DOING WORK IN MOST SCHOOL SUBJECTS.					
55. I AM UGLY.					
56. I LEARNED TO READ EARLIER THAN MOST OTHERS.					
57. I'M GOOD AT THINGS LIKE SPORT, GYM & DANCE.					
58. I HAVE LOTS OF FRIENDS OF THE OPPOSITE SEX.					

	FALSE	MOSTLY FALSE	MORE FALSE THAN TRUE	MORE TRUE THAN FALSE	MOSTLY TRUE	TRUE
59. I SOMETIMES TELL LIES TO STAY OUT OF TROUBLE.						
60. I GET ALONG WELL WITH MY PARENTS.						
61. OVERALL, I'M A FAILURE.						
62. I NEVER WANT TO TAKE ANOTHER MATHEMATICS COURSE.						
63. I DO NOT GET ALONG VERY WELL WITH GIRLS.						
64. I WORRY ABOUT A LOT OF THINGS.						
65. I DO WELL IN TESTS IN MOST SCHOOL SUBJECTS.						
66. I HATE THE WAY I LOOK.						
67. I HATE READING.						
68. I AM AWKWARD AT THINGS LIKE SPORT, GYM, & DANCE.						
69. I GET A LOT OF ATTENTION FROM MEMBERS OF THE OPPOSITE SEX.						
70. CHEATING ON A TEST IS OK IF I DO NOT GET CAUGHT.						
71. I DO NOT LIKE MY PARENTS VERY MUCH.						
72. I AM A USEFUL PERSON TO HAVE AROUND.						
73. I GET GOOD MARKS IN MATHEMATICS.						
74. I MAKE FRIENDS EASILY WITH GIRLS.						
75. I AM A NERVOUS PERSON.						
76. I'M GOOD AT MOST SCHOOL SUBJECTS.						
77. MOST OF MY FRIENDS ARE BETTER LOOKING THAN I AM.						
78. I'M HOPELESS IN ENGLISH CLASSES.						
79. I'M BETTER THAN MOST OF MY FRIENDS AT THINGS LIKE SPORTS, GYM & DANCE.						
80. I'M NOT VERY POPULAR WITH MEMBERS OF THE OPPOSITE SEX.						
81. WHEN I MAKE A PROMISE I KEEP IT.						
82. I HAVE A LOT OF ARGUMENTS WITH MY PARENTS.						
83. I DON'T HAVE MUCH TO BE PROUD OF.						
84. I HAVE ALWAYS DONE WELL IN MATHEMATICS.						
85. I HAVE A LOT IN COMMON WITH THE GIRLS I KNOW.						
86. I OFTEN FEEL GUILTY.						
87. I'M NOT VERY INTERESTED IN ANY SCHOOL SUBJECTS.						
88. I'M BETTER LOOKING THAN MOST OF MY FRIENDS.						
89. I OFTEN HAVE TO READ THINGS SEVERAL TIMES BEFORE I REALLY UNDERSTAND THEM.						
90. I CAN RUN A LONG WAY WITHOUT STOPPING.						
91. MOST BOYS TRY TO AVOID ME.						
92. I SOMETIMES CHEAT.						
93. MY PARENTS ARE USUALLY UNHAPPY OR DISAPPOINTED WITH WHAT I DO.						
94. IN GENERAL I LIKE BEING THE WAY I AM.						
95. I HAVE TROUBLE UNDERSTANDING ANYTHING WITH MATHEMATICS IN IT.						
96. I HAVE FEWER FRIENDS OF THE SAME SEX THAN MOST PEOPLE.						
97. I AM USUALLY RELAXED.						
98. PEOPLE COME TO ME FOR HELP IN MOST SCHOOL SUBJECTS.						
99. NOBODY THINKS THAT I'M GOOD LOOKING.						
100. I LEARN THINGS QUICKLY IN ENGLISH CLASSES.						
101. I AM LAZY WHEN IT COMES TO SPORTS & HARD PHYSICAL EXERCISE.						
102. I HAVE A LOT IN COMMON WITH THE BOYS I KNOW.						
103. I AM HONEST.						
104. IT IS DIFFICULT FOR ME TO TALK TO MY PARENTS.						
105. I CAN DO THINGS AS WELL AS MOST OTHER PEOPLE.						
106. I ENJOY STUDYING FOR MATHEMATICS.						
107. GIRLS FIND ME BORING.						
108. I GET UPSET EASILY.						
109. I'M TOO STUPID AT SCHOOL TO GET INTO A UNIVERSITY.						
110. I HAVE A GOOD LOOKING BODY.						
111. I HAVE TROUBLE TRYING TO EXPRESS MYSELF WHEN I TRY TO WRITE SOMETHING.						
112. I MAKE FRIENDS EASILY WITH MEMBERS OF MY OWN SEX.						
113. I DO NOT GET ALONG VERY WELL WITH BOYS.						
114. IF I REALLY TRY I CAN DO ALMOST ANYTHING I WANT TO DO.						
115. I AM NOT VERY GOOD AT READING.						

	FALSE	MOSTLY FALSE	MORE FALSE THAN TRUE	MORE TRUE THAN FALSE	MOSTLY TRUE	TRUE
116. OVERALL, I HAVE A LOT TO BE PROUD OF.						
117. I AM CHEERFUL AND ON TOP OF THINGS MOST OF THE TIME.						
118. I ENJOY SPENDING TIME WITH MY FRIENDS OF THE SAME SEX.						
119. I FEEL THAT MY LIFE IS NOT VERY USEFUL.						
120. I HAVE TROUBLE WITH MOST SCHOOL SUBJECTS.						
121. I HAVE FEW FRIENDS OF OF THE SAME SEX AS MYSELF.						
122. I DO BADLY ON TESTS THAT NEED A LOT OF READING ABILITY.						
123. I AM A HAPPY PERSON.						
124. BOYS LIKE ME.						
125. MOST THINGS I DO I DO WELL.						
126. I HAVE GOOD FRIENDS WHO ARE MEMBERS OF MY OWN SEX.						
127. OVERALL, MOST THINGS I DO TURN OUT WELL.						
128. NOT MANY PEOPLE OF MY OWN SEX LIKE ME.						
129. MOST GIRLS WANT ME TO BE THEIR FRIEND.						
130. I DON'T GET UPSET VERY EASILY.						
131. NOTHING I DO EVER SEEMS TO WORK OUT RIGHT.						
132. BOYS OFTEN MAKE FUN OF ME.						
133. I GET BAD MARKS IN MOST SCHOOL SUBJECTS.						

	FALSE	MOSTLY FALSE	MORE FALSE THAN TRUE	MORE TRUE THAN FALSE	MOSTLY TRUE	TRUE
134. I SPEND A LOT OF TIME WITH MEMBERS OF MY OWN SEX.						
135. I WORRY MORE THAN I NEED TO.						
136. I MAKE FRIENDS EASILY WITH BOYS.						
137. I AM GOOD AT EXPRESSING MYSELF.						
138. OTHER PEOPLE GET MORE UPSET ABOUT THINGS THAN I DO.						
139. MOST GIRLS LIKE ME.						
140. IT IS DIFFICULT TO MAKE FRIENDS WITH MEMBERS OF MY OWN SEX.						
141. I INTEND TO COMPLETE YEAR 12.						
142. IT'S IMPORTANT TO ME TO BE GOOD AT THINGS LIKE SPORTS, PHYS. ED., GYM, ETC.						
143. IT'S IMPORTANT TO ME TO BE GOOD LOOKING.						
144. IT'S IMPORTANT TO ME TO HAVE A LOT OF FRIENDS OF MY OWN SEX.						
145. IT'S IMPORTANT TO ME TO BE POPULAR WITH MEMBERS OF THE OPPOSITE SEX.						
146. IT'S IMPORTANT TO ME TO DO WELL IN MOST SCHOOL SUBJECTS.						
147. IT'S IMPORTANT TO ME TO DO WELL IN MATHEMATICS CLASSES.						
148. IT'S IMPORTANT TO ME TO DO WELL IN ENGLISH CLASSES.						
149. I INTEND TO GO TO UNIVERSITY AFTER I LEAVE SCHOOL.						
150. IT'S MORE IMPORTANT TO ME TO BE POPULAR WITH SAME-SEX FRIENDS THAN OPPOSITE-SEX FRIENDS.						

NOW WE WANT YOU TO DO A DIFFERENT TASK. Below is a list of personality characteristics. Please use these characteristics to describe YOURSELF. Indicate on a scale from 1 to 7 how true of YOU these various characteristics are. Please do not leave any blanks. As an example consider the characteristic HAPPY. Your answer would be:

1 if it is NEVER OR ALMOST NEVER TRUE that you are happy.
2 if it is USUALLY NOT TRUE that you are happy.
3 if it is SOMETIMES BUT INFREQUENTLY TRUE that you are happy.
4 if it is OCCASIONALLY TRUE that you are happy
5 if it is OFTEN TRUE that you are happy.
6 if it is USUALLY TRUE that you are happy.
7 if it is ALWAYS OR ALMOST ALWAYS TRUE that you are happy.

Thus, if you feel it is SOMETIMES BUT INFREQUENTLY TRUE that you are happy, you should write a "3" next to it: 3 HAPPY

1 NEVER OR ALMOST NEVER TRUE	2 USUALLY NOT TRUE	3 SOMETIMES BUT INFREQUENTLY TRUE	4 OCCASIONALLY TRUE	5 OFTEN	6 USUALLY TRUE	7 ALWAYS OR ALMOST ALWAYS TRUE
___FIRM	___NERVOUS	___WEAK	___LOYAL	___PLEASURE-SEEKING	___DETERMINED	
___DEPENDENT	___AGGRESSIVE	___BASHFUL	___STRONG	___LOVES CHILDREN	___HASTY	
___PATIENT	___CONFIDENT	___MISCHIEVOUS	___CAREFREE	___NEEDS APPROVAL	___BRAVE	
___TENSE	___COMPETITIVE	___RESPONSIBLE	___ABSENT-MINDED	___SENSITIVE TO THE NEEDS OF OTHERS	___LOUD	
___BOSSY	___CASUAL	___EMOTIONAL	___RUDE	___SELF-SUFFICIENT	___LIVELY	
___NOISY	___TIMID	___RESOURCEFUL	___SEES SELF RUNNING SHOW	___SELF-CRITICAL	___CRIES EASILY	
___RASH	___LOGICAL	___SHY	___OUTSPOKEN	___CLEAR-THINKING	___INEFFICIENT	
___SHOW-OFF	___GRATEFUL	___CHILDLIKE	___WORRYING	___SKILLED IN BUSINESS	___HELPFUL	
___INTERESTING	___SARCASTIC	___ANXIOUS	___GENTLE	___FEELS SUPERIOR	___FLASHY	
___APPRECIATIVE	___FORCEFUL	___SOASTFUL	___SILLY	___DEVOTES SELF TO OTHERS	___WIDE INTERESTS	

SSA SURVEY OF SCHOOL ATTITUDES

Intermediate Answer Sheet
(Not machine scorable, Non Picture, UCR format.)

Form A [ ]

M R C S

Name_____Grade____ Male [ ] Female [ ]

Teacher_____School_____

City_____State ____ Today's Date _____

Sample A	Like	Not Sure Don't Care	Dislike	Sample B	Like	Not Sure Don't Care	Dislike
Eating ice cream cones	✓	( )	( )	Washing the floor	( )	( )	( )

		Like	Not Sure Don't Care	Dislike
1.	Learning what numbers mean	( )	( )	( )
2.	Finding out how different people live	( )	( )	( )
3.	Seeing how plants grow	( )	( )	( )
4.	Writing letters to friends	( )	( )	( )
5.	Adding numbers	( )	( )	( )
6.	Learning about cities	( )	( )	( )
7.	Studying about the stars	( )	( )	( )
8.	Reading to the class	( )	( )	( )
9.	Telling which number comes next in the series	( )	( )	( )
10.	Learning about our country	( )	( )	( )
11.	Finding out what different animals eat	( )	( )	( )
12.	Making words that sound alike	( )	( )	( )
13.	Multiplying big numbers	( )	( )	( )
14.	Finding out about the President's job	( )	( )	( )
15.	Learning about rock formations under the ground	( )	( )	( )
16.	Reading stories	( )	( )	( )
17.	Finding the missing number	( )	( )	( )

SSA - Page 2

		Like	Not Sure Don't Like	Dislike
18.	Learning about different parts of a city from a map	( )	( )	( )
19.	Working with magnets	( )	( )	( )
20.	Finding words that have the same meaning	( )	( )	( )
21.	Doing hard subtraction problems	( )	( )	( )
22.	Reading books about history	( )	( )	( )
23.	Learning about different parts of the body	( )	( )	( )
24.	Listening to poems	( )	( )	( )
25.	Measuring things	( )	( )	( )
26.	Learning about different ways to travel	( )	( )	( )
27.	Finding out how motors work	( )	( )	( )
28.	Telling stories to the class	( )	( )	( )
29.	Dividing numbers	( )	( )	( )
30.	Studying the weather in different places	( )	( )	( )
31.	Learning how propellers move ships	( )	( )	( )
32.	Using the encyclopedia	( )	( )	( )
33.	Solving word problems	( )	( )	( )
34.	Learning what causes pollution	( )	( )	( )
35.	Studying about things that live in the water	( )	( )	( )
36.	Learning to spell	( )	( )	( )
37.	Finding out about different kinds of numbers	( )	( )	( )
38.	Learning about the kinds of buildings made by different people	( )	( )	( )
39.	Finding out what causes storms	( )	( )	( )
40.	Getting a book as a gift	( )	( )	( )
41.	Learning about fractions	( )	( )	( )
42.	Finding out how courts and judges work	( )	( )	( )
43.	Learning about solids, liquids, and gases	( )	( )	( )

SSA - Page 3

		Like	Not Sure Don't Like	Dislike
44.	Reading poems	( )	( )	( )
45.	Multiplying numbers in different ways	( )	( )	( )
46.	Learning about different groups of people	( )	( )	( )
47.	Finding out why some things float	( )	( )	( )
48.	Making up a title for a story	( )	( )	( )
49.	Learning about weights	( )	( )	( )
50.	Studying about famous people	( )	( )	( )
51.	Learning about insects	( )	( )	( )
52.	Going to the library	( )	( )	( )
53.	Learning how to use charts	( )	( )	( )
54.	Using maps	( )	( )	( )
55.	Finding out how the heart works	( )	( )	( )
56.	Writing reports for school	( )	( )	( )
57.	Working with sets	( )	( )	( )
58.	Learning what taxes pay for	( )	( )	( )
59.	Measuring how fast things go	( )	( )	( )
60.	Choosing a good word for a sentence	( )	( )	( )

STOP