DOCUMENT RESUME

ED 378 240 TM 022 618

AUTHOR Winfield, Linda F.
TITLE Dimensions of the Problem of Testing and Assessment of At-Risk Students.
INSTITUTION Research for Better Schools, Inc., Philadelphia, Pa.
PUB DATE Nov 87
NOTE 21p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Change; Educational Policy; Educational Testing; Elementary Secondary Education; *Equal Education; *High Risk Students; Literacy; *Social Influences; Test Bias; *Test Use
IDENTIFIERS *Instructional Sensitivity; *Reform Efforts

ABSTRACT

Implications of the educational reform movement's emphasis on testing and assessment are considered as they touch on at-risk students. The four dimensions that must be considered in addressing the assessment of at-risk student populations are (1) antecedent instructional conditions; (2) type of test or assessment; (3) social and educational demands; and (4) educational policy decisions based on test results. These dimensions are interrelated, as examples from real-life test use illustrate. They must be considered in the context of two critical educational and social demands with significant impact on assessment and education. First is the demand for instructional sensitivity in assessment, and second is the demand for higher levels of literacy in today's society. These demands compete with demands for policy-oriented tests, which may have an adverse impact on educational opportunities available for high-risk students. Every effort must be made to guarantee instructional sensitivity and equity in testing to assure at-risk students the best possible educational outcomes. (Contains 33 references.) (SLD)

DIMENSIONS OF THE PROBLEM OF TESTING AND ASSESSMENT OF AT-RISK STUDENTS

by

Linda F. Winfield, Ph.D.
Assistant Professor
Temple University
College of Education
Philadelphia, PA 19122

Research for Better Schools, Inc.
444 North Third Street
Philadelphia, PA 19123

November 1987

DIMENSIONS OF THE PROBLEM OF TESTING AND ASSESSMENT OF AT-RISK STUDENTS

The recent reform movement in education, with its increased focus on educational outcomes, has resulted in more testing and assessment of students at all educational levels. This paper discusses some dimensions for considering the implications and impact of this movement on at-risk student populations. It consists of four sections. The first section describes four key dimensions of the problem of increased testing and assessment of at-risk students. The second section illustrates the interrelatedness of the dimensions through discussions of major issues. A third section discusses two key educational and social demands which have important implications for the testing and assessment of at-risk students. The final section makes some overall conclusions.

Dimensions of the Problem

The educational accountability movement initiated in the late 1960's continues to gather high public support. Accountability-inspired legislation, such as student competency testing and entrance and exit proficiency exams for high school and college students, is a growing trend. This movement grew out of the belief of many citizens and policy makers that the quality of education is poor and that there is a need to focus attention on the outcomes of educational institutions. The assessment movement is of particular concern to at-risk and low socioeconomic status groups as education is considered to be a key mechanism for improving their status and mobility.

There are four dimensions which must be considered when addressing the problem of testing and assessment of at-risk student populations. They are: antecedent instructional conditions, type of test or assessment,

social and educational demands, and educational policy decisions based on test results. These four dimensions or categories are applicable to all levels of education and all students, e.g., elementary and secondary students and at-risk minority and at-risk non-minority populations. For example, standardized test scores of at-risk students in Appalachia can be compared to the national average (Adams, 1984; Caudill, 1976; Ford, 1962) and the implications for this group can be compared to those of other at-risk groups. The four proposed dimensions for considering implications are discussed below.

The first dimension, antecedent instructional conditions, focuses on a student's prior opportunity to acquire the appropriate content, skill, and knowledge embodied in the test. These conditions reflect factors such as classroom and remedial instruction; practice, feedback and motivation; academic engaged time; high school curriculum track; test-taking skill; academic courses taken at the secondary level; quality of teaching and counseling; parental academic support and encouragement; and out-of school learning opportunities. (See Ogbu, 1978 for a comprehensive societal framework for these conditions).

The second dimension, type of test or assessment, has been a continual object of debate with regard to at-risk populations. There are two major issues: first, how accurately or validly the test measures the learning outcomes of at-risk students; and second, how accurately the test predicts at-risk students' future success. Controversy around these issues has generally focused on aspects of test design and construction, such as item bias, test bias, and test validity and reliability.

The third dimension, social and educational demands, reflects the context and requirements surrounding test development. Historically, the

testing movement began in a period when there was a need for fast, objective methods for identifying persons most likely to succeed in designated situations or treatments (Gordon & Terrell, 1981). Currently, there are two competing social and educational demands. On one hand, most school building administrators and teachers support the need for tests whose results can be used to improve instruction. On the other hand, the educational accountability movement has fueled an increased use of test results for administrative purposes, such as monitoring and evaluating school and program outcomes.

The fourth and last dimension includes educational policy decisions which are based on test results. Much of the discussion of implications of testing for at-risk and low status groups has focused on this dimension. It includes policy decisions such as how to allocate resources and how to place or select students for "special" programs. A major argument here surrounds the use of tests in controlling the opportunity and reward structures which maintain the status quo. Those against using tests in these ways support banning standardized tests.

Within each of these four dimensions or components, the implications of testing for at-risk students reside on a continuum that ranges from optimal to adverse, from constructive to destructive, or from positive to negative. However, it is important to remember that the four dimensions are interrelated, and to avoid the tendency to isolate a single dimension. The following section deals with both the interrelatedness and range of these dimensions in discussing some major issues.

Interrelatedness of Dimensions

There are many examples that indicate both the interrelatedness of the four dimensions or components and the poles of the continuum for each

dimension. For example, the recent controversy surrounding the Florida minimum competency tests centered on the following issues: had at-risk students received the "antecedent instructional conditions" prior to testing, were the tests valid, and were appropriate decisions made as a result of testing.

Another example is the frequent use of the results of standardized achievement tests or IQ tests to classify and/or place at-risk students in special classes for the educable or mildly mentally retarded. Because of the disproportionate number of classifications and placements of minority students in these categories, the use of tests for this purpose is highly controversial and has been discontinued as early as 1975 in California. Although technical issues related to the test itself, such as test bias, are relevant, of equal concern are the social and educational demands and educational policy decisions that result from test use. For example, these tests were criticized as perpetuating negative sterotypes, as diminishing opportunities for minority students, and as being part of a process whereby greater numbers of minority students were classified and placed in ineffective educational programs (Bernal, 1975; Jackson, 1975). The focus here is with social justice regarding the use of test results rather than with the theoretical and technical aspects of testing (Gordon & Terrell, 1981).

Although the use of IQ tests for student classification or placement was banned in California in 1975, this decision had no appreciable effect on the number of Black students in California placed in special education classes until 1979, when court ordered rigid quotas were imposed (Reschly, 1981). Although this author supports the moratorium on IQ testing, banning the test as a solution to minority overrepresentation fails to consider other factors which also contribute to the lower academic performance and

test scores of at-risk students. One key factor is antecedent instructional conditions.

Another example of the interrelatedness of the dimensions, although not directly related to the at-risk student population, involves a recent case concerning teacher competency testing in Alabama. The case centers on the validity of the operational cut scores on a teacher competency test, the Alabama Basic Professional Studies Test. If the cut score selected by the state had been implemented, only one examinee from all of Alabama's institutions would have passed at the first administration. Madaus (1986) argued successfully that the cut score the state selected was not valid, and should be lowered by three standard errors. His position was that, in high stakes tests like the one in Alabama, the issue of content validity is separate from the issue of the degree to which the test accurately distinguishes those who actually have the skill and/or knowledge necessary to teach from those who do not. His argument is based on the belief that, although one can distinguish between a decision and the test used to make the decision, for most high stakes tests the decision is inherently linked to the test. Thus, the validity of the decision is directly related to the validity of the test.

In terms of their implications, the above examples represent the negative end of the continuum. We might also consider a few examples from the positive end. Although most standardized tests provide limited diagnostic information regarding individual student learning outcomes, some schools and districts use aggregate test results in attempts to support specific school or district improvement programs or initiatives. These programs or initiatives may include revising curricula, monitoring student progress, and assessing strengths and weaknesses in curricula coverage

(Fullan, 1985; Purkey & Smith, 1985; Venezky & Winfield, 1979; Winfield, 1987b). This is a program evaluation approach. It attempts to determine whether improvements or declines in student learning are a direct result of components of curriculum and instruction within the school. Thus, the burden of accountability is on the instructional program, the teacher, and the school. This approach is to be differentiated from the exit requirement approach which places the burden of accountability on the student, rather than the school, and penalizes the student, rather than the educational system, for failure (Tuckman, 1985).

If competency-based tests are used in a program evaluation framework, they are likely to have a positive impact on at-risk student achievement, although less positive than their impact on non-at-risk student achievement. Exploratory studies of 1983-84 National Assessment of Educational Progress (NAEP) data compared the performance of a nationally representative sample of Black students in schools with minimum competency testing (MCT) programs to comparable students in schools without such programs (Winfield, 1987a). After the studies adjusted for school-level, socioeconomic status, school racial composition, parental education, region of the country, and student academic behaviors, the average reading proficiency of students in MCT schools was higher at two of the three grade levels examined, i.e., higher at 8th and 11th grades but not at 4th grade. Conclusions based on these studies are tenuous and one can only speculate concerning the cause of the higher reading proficiency scores. The increased scores may not have resulted from the tests, per se, but from some other unmeasured but related school characteristic associated with a competency-based approach, e.g., increased monitoring of student progress, and additional time and resources targeted to at-risk students.

Alternatively, the results at the 11th grade may have been impacted by the dropout rates of at-risk students in schools with MCT programs. Data analyzed in these studies were not sufficient to pinpoint the exact cause of the improved proficiency.

The educational benefits and costs to be derived from competency testing of at-risk students continue to be an open and debated issue. While it may be the case that increased resources are targeted towards students needing remediation, Black students fail the tests in substantially higher proportions than do white students (Jaeger, 1982; Linn, Madaus & Pedulla, 1982; Serow, 1984). Until the impact of MCT requirements is further evaluated, the requirements are likely to continue having an adverse disproportionate impact on Black and other at-risk students (Linn, Madaus & Pedulla, 1982).

The issue of educational benefit and cost is central to the recent educational reform movement, with its increased emphasis on testing. McDill, Natriello & Pallas (1986) note some potential negative outcomes of what may be assumed to be a common educational benefit:

> If academic standards are raised and students are not provided with substantial additional help to attain them, it seems reasonable to expect that at-risk students - those socially and academically disadvantaged - will be more likely than ever to experience frustration and failure. The result for these students may not be notable increases in cognitive achievement, but rather notable increases in absenteeism, truancy, school-related behavior problems and dropping out. (p. 122)

Another key issue to consider in the testing of at-risk student populations is the issue of test validity. The public controversy and demands regarding test validity (Nader, 1987) are useful and necessary to sensitize test publishers and educational institutions to the influence of differences in ethnicity, gender, social class, and culture in the

development and assessment of cognitive performance (Gordon & Terrell, 1981). However, there remain two major concerns. One concern is that very little of this discussion is directed toward and will impact state school boards, departments of education, and/or local districts where decisions concerning the use of competency tests for promotion and graduation are being made. A second concern is that an exclusive focus on psychometric issues and test use and decision making alone will not account for 100 percent of the variance in at-risk student achievement. In order for meaningful change and improvement, there must also be a concerted effect to improve the "antecedent instructional conditions" of at-risk students and low status groups.

Educational and Social Demands

There are two critical educational and social demands which have significant implications for the testing and assessment of at-risk students and which have the potential to shape future debates regarding the testing of at-risk student populations. They are, first, the demand for instructionally-sensitive assessments, and second, the demand for higher levels of literacy in today's society. These demands are discussed below.

Demand for Instructional Sensitivity

In a discussion of access to education, Glaser (1981) indicates that historically, the use of tests for student selection and screening, and for predicting student performance originated during a period when school attendance was made compulsory and educators needed ways to deal with a wide range of student diversity. However, this type of system is no longer a current educational demand. Accordingly to Glaser, there is now less emphasis on selecting individuals for available educational opportunities

and more on helping them to succeed in these opportunities. It is interesting to note that for Blacks and other minorities, tests were used historically to restrict rather than select educational opportunities (Sewell, 1987). Glaser's point, which is well taken, is that tests have not traditionally been designed to guide specific instructional practices. Yet teachers and schools need information on their students for instructional decision making rather than prediction. Glaser (1986) notes:

> At present, tests (with the exception of the important informal assessments of the good classroom teacher) typically are not designed to guide the specifics of instruction. We use them primarily as indicators to signal general rises or declines in school performance. They serve as an index to the standards of schools, but they are not designed to shape progress effectively toward these standards and can do so only indirectly, if at all. In the 21st century, tests and other forms of assessment will be valued for their ability to facilitate constructive adaptations of educational programs.
>
> To accomplish this, students and teachers will need information that can inform instructional decisions rather than just predict academic success or offer a percentile or grade-level index of relative standing and global attainment. The information required will be analogous to that used by an opera teacher or a swimming coach to guide the development of further competency and proficiency. Testing and teaching will be integral events. (p. 46)

Gordon & Terrell (1981) provide a similar view. They describe testing in education as being influenced by the meritocratic approach, an approach which identifies individuals meriting special attention. They argue that this approach, with its emphasis on sorting, predicting, and selecting, influenced the development and use of psychometric tests. Furthermore, these tests were only appropriate in the past, when resources for the development and nurturance of human potential were quite scarce. Gordon & Terrell (1981) describe this changing view of assessment as follows:

> In the last three decades, a commitment to democratic access to human developmental resources and opportunities has been repeatedly voiced; indeed, recent history has been marked by

> the assertion of a constitutional right to equal opportunity in education and improvement. Under such a commitment the purpose of assessment can no longer be to sort, to predict, and to select with a view to identifying and rewarding a chosen few who are most likely to succeed. Use of tests that were developed to meet the assessment needs of the early 20th century has drawn fire because it is neither in keeping with the changing political and social context nor with the requirements of assessment in this new context. (p. 1168)

This educational demand to abandon the meritocratic approach will most likely have a positive impact on at-risk students. It gives priority to the evaluation of antecedent instructional conditions and to assessment measures which are sensitive to those conditions. This new approach will not merely categorize individuals in terms of cognitive functioning, but also describe the processes by which learning is facilitated or inhibited so that effective educational intervention strategies can be developed (Gordon & Terrell, 1981; Sewell, 1987).

Despite the optimism of cognitive psychologists and psychometricians regarding the potential for future improvements in instructionally-oriented tests, there has been a dramatic increase in policy-oriented testing during the last decade (Airasian, 1987). The functions of these policy-oriented tests have been to monitor, manage, and control rather than to improve instructional outcomes. For example, the increased administrative use of tests and test results for decisions involving high school graduation and remedial funding makes the test a significant mechanism for distributing and redistributing resources. Airasian (1987) notes:

> As long as educational benefits and resources are considered important for success in our society, and as long as these benefits and resources are comparatively scarce, the mechanism used to distribute them will be extremely important to individuals and groups who seek the benefits and resources. In essence, test scores become a medium of exchange to be bartered for educational, social, and economic benefits or rewards. (p. 409)

Thus, in discussing the implications and impact of testing on at-risk student populations, a distinction must be made between two types of testing; first, instructionally-oriented testing that is used to assess learning outcomes, and second, policy-oriented testing that is used to assess the effect of an educational policy or program. Airasian (1987) elaborates on this distinction using the element of control. He describes instructionally-oriented testing as being controlled by local school districts as compared to policy-oriented testing which is controlled by external agencies.

Demand for Increased Literacy

A second related educational and social demand which has significant implications for the testing and assessment of at-risk students is the demand for high levels of literacy in today's society. In an era of rapid technological change, new developments in micro-electronics, biological science, and other high tech areas are transforming employment needs and opportunities in virtually all sectors of the economy. Although this transformation may provide renewed economic prosperity for our nation, it will most likely profoundly alter the nature of work in our society (Rumberger & Levin, 1984). From their studies of literacy in various cultures, Scribner and Cole (1981) conclude that:

> As the technology of a society becomes more complex and it becomes more closely integrated into world affairs, we can expect the number and variety of literacy practices to increase, bringing with them new skills or more complicated versions of old skills. (p. 258)

Many of these new or revised skills will be higher order thinking skills, such as problem-solving and analysis, which go far beyond literal level skills (Mikulecky, 1982).

Increased literacy skills for at-risk and minority youth are critical to their successfully achieving occupational status, and more importantly, for converting this status into increased earnings (Greenfield, 1980). Although demographic projections indicate that by 1990 the total number of new labor force entrants aged 18 to 24 years old will decline, at-risk youth will comprise a greater proportion of all entry level workers (National Council on Employment Policy, 1984). According to Venezky, Kaestle and Sum (1987):

> Unless improvements are made in American education, the labor force will become progressively less capable of doing highly skilled work as the proportion of poorly educated, low-income workers increases throughout the remainder of this century. Perhaps more important, the present inequalities, if not ameliorated, could lead to a society more deeply divided along racial and socioeconomic lines, with the number of low-income, low-achieving persons increasing from year-to-year. (p. 3)

Within the context of changing demographics and technology, there is reason for concern over the inadequate literacy skills of America's at-risk and minority youth. Data from two recent national assessments[1] support this concern. Results from the NAEP Reading and Writing Assessment (see National Assessment of Educational Progress, 1985) indicate that although the average performance of at-risk groups improved, the achievement gap between racial/ethnic groups continues to exist. The data specifically show notable gains for Black and Hispanic students since their first assessment in 1971. While white students made modest improvements at all ages, Black students showed an increase in performance at age 17, continued improvement at age 13, and a leveling off of performance at age 9. White students also improved significantly at age 17, but not at ages 9 or 13. Hispanic students, in contrast, showed continuing, relatively steady improvement at all three age levels.

In interpreting these and similar results, Carroll (1987) notes that focusing on a group's average performance is misleading, as one generally does not know in which part of the distribution any changes (i.e., increases or decreases) have taken place. Using the 1984 NAEP data, if performance is examined across levels of the scale, only 16 percent of Black 17-year olds demonstrated "adept" reading skills and strategies compared to 45 percent of the white students at this age level. Performance at the adept level indicates the ability to find, understand, summarize, and explain relatively complicated information. These figures are for an in-school population, and when one considers differential dropout rates by race, the literacy problem of Black youths is compounded.

The results of the NAEP Young Adult Literacy Assessment (Kirsch & Jungeblut, 1986) indicate the overwhelming majority of young adults adequately perform lower level tasks, e.g., filling out a job application and locating "net pay" on a wage and tax statement. However, a sizable number appears unable to do well on tasks of moderate complexity, and only a relatively small percentage of this group is estimated to be able to perform at levels typified by more complex and challenging tasks, e.g., writing a letter to explain an error in a monthly credit-card bill. The fact that fewer and fewer individuals attain these moderate and high levels of proficiency is most pronounced for young adults who terminated their education early and for at-risk group members (National Assessment of Educational Progress, 1986). Only 31 percent of Black young adults were estimated to be at or above the average reading proficiency of 11th graders as measured in the 1985 NAEP in-school assessment; 53 percent were estimated to be at or above the average 8th grade proficiency; and 82 percent

were estimated to be at or above the average reading proficiency of 4th graders.

According to the 1986 NAEP data, Black young adults, on the average, performed significantly below white young adults--by almost a full standard deviation. Hispanic young adults, on the average, performed about midway between their Black and white peers. This difference appeared at each educational level. However, there was a strong relationship between literacy skills and formal educational attainment for the entire sample as well as for each racial/ethnic subgroup. As might be expected, Black and Hispanic high school graduates who did not go on to college performed less well than those who were four-year college graduates. Venezky, Kaestle, and Sum (1987) report:

> The poor performances of many black and Hispanic high school graduates who did not go on to college should be viewed as a major national concern by educators and economic policy-makers. During the past eight years, the proportion of black high school graduates (18-24) enrolled in college has been declining in both absolute terms and relative to white high school graduates. (p. 35)

These authors further suggest that, although economic factors contribute to this decline, e.g., rising costs of education and declining income of Black adult males, the limited academic achievement of many Black high school graduates negatively influences college enrollment decisions and continues to contribute to the widening of the gap in Black/white economic status (Harrington, Sum, & Fogg, 1986).

Conclusion

The need for instructionally-oriented tests and higher levels of competence are two educational and social demands that will potentially reshape the debates on the testing and assessment of at-risk student populations. These demands compete, however, with demands for policy-

oriented tests which may have an adverse impact on the educational opportunities available to at-risk students. In considering the impact of tests on this special population, continual efforts must be exerted to: (1) improve antecedent instructional conditions, (2) evaluate the actual tests in terms of equity considerations, (3) demand more instructionally-sensitive measures, and (4) monitor and scrutinize policies and decisions which are based on test results of at-risk students.

NOTES

[1]The purpose of the national assessment is to provide the education community and the general public a view of the trends and state of educational progress of America's school-age and young adult populations. NAEP has been in existence since 1969 and has involved a random cross sectional survey of in-school 9, 13, and 17-year olds. NAEP is funded by the federal government. It is currently under contract to the Educational Testing Service.

In the 1983-84 assessment of reading and writing, in addition to sampling by age, students of the corresponding model grades 4, 8, and 11 were also sampled. Each age/grade cohort included approximately 30,000 students. The NAEP sample was based on a highly stratified, 3-stage sampling design in which first counties, second schools, and third students were sampled.

In 1985, NAEP assessed the literacy skills of America's young adults. NAEP used a wide variety of tasks that simulated the diversity of literacy activities that people encounter at work, at home, at school, and in their communities. To conduct this assessment, NAEP drew a nationally representative sample of 3,600 young adults aged 21 to 25 years. Blacks and Hispanics were sampled at twice the rate of whites. In-depth, household interviews were conducted, which included the collection of background information as well as the assessment of everyday literacy tasks.

REFERENCES

Adams, G. (1984). People of the Appalachians. (ERIC Document Reproduction Service No. ED 242 651)

Airasian, P. (1987). State mandated testing and educational reform: Context and consequences. American Journal of Education, 95(3), 393-412.

Bernal, E. M., Jr. (1975). A response to educational uses of testing with disadvantages subjects. American Psychologist, 30, 93-95.

Carroll, J. B. (1987). The national assessment in reading: Are we misreading the findings. Phi Dela Kappan, 68(6), 424-430.

Caudill, H. M. (1976). The watches of the night. Boston, MA: Little, Brown & Co.

Ford, T. R. (Ed.)(1962). The Southern Appalachian Region. Lexington, KY: University of Kentucky Press.

Fullan, M. (1985). Change processes and strategies at the local level. The Elementary School Journal, 85(3), 391-422.

Glaser, R. (1986). The integration of instruction and testing. In E.E. Freeman (Ed.), The Redesign of Testing for the 21st Century: Proceedings of the 1985 ETS Invitational Conference. Princeton, NJ: Educational Testing Service.

Gordon, E., & Terrell, M. D. (1981). The changed social context of testing. American Psychologist, 36(10), 1167-1171.

Greenfield, S. (1980). The human capital model and american youths: The roles of schooling experience and functional literacy. Austin, TX: Southwest Educational Development Laboratory. (ERIC Document Reproduction Service No. ED 195 772)

Harrington, P., Sum, A., & Fogg, W. (1986). Declining real incomes of adult men in the US 1973-1985: Implications for education and training programs. Boston: Northeastern University, Center for Labor Market Studies.

Jackson, G. (1975). On the report of the ad hoc committee on educational uses of tests with disadvantaged students. American Psychologist, 30, 88-92.

Jaeger, R. M. (1982). The final hurdle: Minimum competency achievement testing. In G. R. Austin & H. Garber (Eds.), The rise and fall of national test scores (p. 223-246). New York: Academic Press.

Kirsch, I. S., & Jungeblut, A. (1986). Literacy: Profiles of America's young adults. Princeton, NJ: National Assessment of Educational Progress.

Linn, R. L., Madaus, G. F., & Pedulla, J. (1982). Minimum competency testing: Caution on the state of the art. American Journal of Education, 91, 15-19.

Madaus, G. (1986). Measurement specialists: Testing the faith - a reply to Mehrens. Educational Measurement Issues & Practices, 5(4), 11-16.

McDill, E. L., Natriello, G., & Pallas, A. M. (1986). A population at risk: Potential consequences of tougher school standards for student dropouts. American Journal of Education, 94(2), 135-181.

Mikulecky, L. J. (1982). Job Literacy: The relationship between school preparation and workplace actuality. Reading Research Quarterly, 17(3), 400-419.

Nader, R. (1987). Sixty years of idiocy is enough. The Fair Test Examiner. Cambridge, MA: National Center for Fair & Open Testing.

National Assessment of Educational Progress. (1985). The reading report card: Progress toward excellence in our schools (Report No. 15-R-01). Princeton, NJ: National Assessment of Educational Progress.

National Council on Employment Policy (1984). Investing in America's future. Alexandria, VA: Remediation and Training Institute.

Ogbu, J. (1978). Minority education and caste. New York: Academic Press.

Purkey, S. C., & Smith, M. S. (1985). School reform: The district policy implications of the effective schools literature. The Elementary School Journal, 85(3), 353-390.

Reschly, D. J. (1981). Psychological testing in educational classification and placement. American Psychologist, 36(10), 1094-1102.

Rumberger, R. W., & Levin, H. M. (1984). Forecasting the impact of new technologies on the future job market (Project Report No. 84-A4). Stanford University Institute for Research on Educational Finance and Governance.

Scribner, S., & Cole, M. (1981). The psychology of literacy. Cambridge, MA: Harvard University Press.

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

Sewell, T. E. (1987, November). Intellectual assessment of at-risk students: Classification vs. instructional goals. Invited address at the Cross-Laboratory Conference on Teaching Thinking and At-Risk Students, Philadelphia, PA.

Tuckman, B. W. (1985). Evaluating instructional programs. Boston: Allyn & Bacon.

Venezky, R. L., Kaestle, C., & Sum, A. (1987). The subtle danger. Princeton, NJ: Center for the Assessment of Educational Progress.

Venezky, R. L., & Winfield, L. F. (1979). Schools that succeed beyond expectations in teaching reading (Technical Report #1). Newark, DE: University of Delaware Studies on Education:

Winfield, L. F. (1987a). The relationship between minimum competency testing programs and students' reading proficiency: Implications from the National Assessment of Educational Progress 1983-84 (Research Report #87-15). Princeton, NJ: Educational Testing Service.

Winfield, L. F. (1987b). Teachers' estimates of test content covered in class and first grade students' reading achievement. Elementary School Journal, 87(4), 437-454.