

Tutoring Using the PACE Learning Systems to Remediate Students who Fail High School Basic Skills Exams

Tim Lull, Superintendent
Walton County Schools, GA, U.S.A.

James L. Pate, Assistant Professor
Department of Educational Leadership
Valdosta State University, GA, U.S.A.
jl pate@valdosta.edu

Nicole Morgan Gibson, Assistant Professor
Department of Educational Leadership
Valdosta State University, GA, U.S.A.
nmgibson@valdosta.edu

John Freeman, Assistant Professor
Department of Educational Leadership, Policy and Technology Studies
University of Alabama, AL, U.S.A.

Abstract: Many states mandate that students pass a standardized test with a minimum score for either promotion or graduation. This study looked at students who were not successful in passing the standardized test on the initial attempt and the effect of remediation on their subsequent attempts. Variables used in this study included race and gender. Although the study was comprised of a limited sample, the results indicated a positive link between the remediation treatment and success on subsequent attempts to pass the standardized test.

The close of the twentieth century brought with it unprecedented scrutiny of the failure of American educational systems to adequately teach the nation's children. In 1983, the US Department of Education's National Commission on Excellence in Education published *A Nation at Risk*, a report that concluded, "If an unfriendly power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war. As it stands, we have allowed this to happen to ourselves" (p. 5). The national response to the state of American education resulted in unparalleled accountability measures set in place by state boards of education throughout the nation (Airasian, 1988; Airasian, 1993; Madaus, 1991).

Today we are faced with another federal mandate, the *No Child Left Behind Act of 2001 (NCLB)*. Although *NCLB* touches on many areas of our educational system, the testing and accountability measures of student achievement required of state and local school districts will likely have the most impact on education (Kucerik, 2002). Under this law, states are charged with the development of challenging content standards in reading and mathematics, a system of annual testing in grades three through eight to determine whether academic achievement standards are being met by the state, local districts, and individual schools, and the alignment of the state's academic content and academic achievement standards (U.S. Department of Education, 2002).

The primary function of assessing student progress through standardized tests has given way to allowing the results of these tests to determine promotion, retention,

and graduation for students, as well as continued employment for educators (Airasian, 1993). More than 25 states, including Georgia and Alabama, turned to standardized testing in order to develop a uniform measure to rank the status of school districts, schools, teachers, and students as compared nationally and internationally prior to *NCLB*. The requirements established under this Act have made these testing programs more significant for educators as the results are now required to be publicized extensively and districts are required to demonstrate progress toward minimum levels of proficiency within twelve years (U.S. Department of Education, 2002). Additionally, many states, including Georgia, require students to pass a standardized graduation exam in order to graduate from high school. As a result, school districts in Georgia searching for ways to ensure their students are prepared to meet this standardized testing requirement for both the Georgia Criterion Competency Test and the Georgia Graduation Test.

Many of these reforms with a focus on testing and higher academic standards are the precursors of today's controversial accountability movement. Surprisingly, there has been very little study of the consequences of the movement despite the considerable discussions of the pros and cons of this movement (Dee, 2002). There is support for arguments on both sides of the issue. Some proponents of minimum competency, high-stakes tests claim that when important consequences are attached to individual performance on achievement exams, students and teachers are motivated to succeed (Manzo, 1997). Others assert that testing and accountability systems will improve classroom instruction arguing that tests can identify problems that would otherwise go undetected and informed parents will be empowered to ensure that schools provide a quality education (Kucerik, 2002). Opponents argue that students should not be judged on the basis of one test, especially when their academic futures are at stake (Manzo, 1997). They also argue that schools will teach to the test and redirect their resources away from teaching critical thinking skills to exclusively drilling the material tested (Kucerik, 2002). Others argue that testing bias is an issue to be considered when discussing the high stakes testing movement. This issue will be addressed below. Despite the controversies, policies that support standardized testing continue to enjoy widespread public support. They continue to be seen by many as a way to raise academic standards, hold educators accountable for meeting the standards, and increase public confidence in schools (Manzo, 1997; Kucerik, 2002).

By using standardized tests as the indicator of schools' success or failure, states are ratcheting the stakes higher and creating programs that tie the test results to graduation. One clearly definable group that suffers disproportionately from the consequences of high-stakes testing is students who fall through the cracks and for whom there is no safety net. Researchers identify those students most likely to be at-risk of failing academically as African-American, Hispanic, and/or low-socioeconomic status (Archer & Dresden, 1986; Boone, Rogg, Kahle, & Damnjanovic, 1997; Griffin & Heidorn, 1996; Hampton, 1997; Ivory, 1993; Kohn, 2002; Kucerik, 2002; Lam, 1995; Lawton, 1997; Lindsay, 1994; Manzo, 1997; Samuelsen, 2001; Vanneman, 1998a, 1998b, 1998c). Researchers often analyze educational programs, test results, and educational and economic opportunities to determine whether bias exists on the basis of race/ethnicity, socioeconomic status,

and gender. In studies done by Bruschi and Anderson (1994) on achievement in science, differences in scores existed between gender and racial groups. Not only did they find differences based on these variables, they determined that gender differences became more significant with age. Presently, racial and ethnic minorities, who are often also socioeconomically disadvantaged, comprise a disproportionate number of students in remedial and lower-track courses because of their poor performance on traditional assessments (US Department of Education, 1996).

The American Educational Research Association (AERA) recently acknowledged that policy makers generally institute high stakes testing with good intentions of improving education, but they need to carefully evaluate the tests' potential to cause serious harm. AERA recommends that educators be cautious about reaching decisions that affect an individual's educational opportunities based on the score of one test alone. They also recommend alignment of curriculum to the test, varied and meaningful opportunities for students to learn the content, teacher retraining to support student learning of the content, meaningful remediation opportunities for students, and multiple chances to pass the assessment if the scores are used to determine promotion or graduation (Education World, 2000).

Intervention in the educational process is necessary to remediate students for whom the traditional school environment provides an inadequate means for preparing them to pass these graduation and other standardized tests. Research on tutorial remedial programs designed to help students pass minimum competency tests is limited. Nevertheless, noteworthy findings indicate that, on the second competency test, students who participate in remediation programs outperform those who are not given remediation; however, remediation does not seem to effectively overcome deficiencies for those who repeatedly fail minimum competency tests (Abbott, Connell, Davidson, & McLean, 1982).

Graduation requirements have been gradually increased and standardized tests have been developed to determine student mastery of required skills (Ingram, 1995). As the difficulty of graduation tests increase, so do the number of students who fail one or more portions of these exams. In schools where students do not make passing scores on these tests, states require them to develop remediation programs for students who score below the standards set by state and federal education agencies.

Purpose of the Study

Standardized testing of American students is regarded as one method to assess academic progress of these students and to compare American students with their foreign counterparts (National Commission on Excellence in Education, 1983). Today, high-stakes testing is the primary tool used to judge students, teachers, administrators, and school systems (Airasian, 1993). Children are tested to an extent unprecedented in our history. While American children have had to take tests previously, never have tests been given so frequently, nor have they played such a prominent role in schooling (Kohn, 2002).

The results of the Alabama *High School Basic Skills Exit Exam*, the Georgia *High School Graduation Test*, and similar tests in other states indicated that traditional methods of instruction were not successfully reaching all of the students. New and innovative approaches needed to be explored and implemented. This study

will focus on efforts made in the state of Alabama. It is hoped that the results from this study can be applied to Georgia and other states with similar testing and remediation requirements. While new strategies of instruction were initiated across the curriculum, remediation classes were developed for those students who had failed one or more sections of a high school basic skills test before the end of their high school careers. The goal of the program was to successfully remediate skills so those students who had previously failed one or more portions of a high school basic skills exit exam would pass the test upon reexamination. Therefore, the purpose of this study was to compare the Alabama *High School Basic Skills Exit Exam* passing rates of students who participated in tutorial remediation using the PACE Learning Systems program with a comparable group of students who chose not to participate. In addition, gender and ethnicity differences were explored for the two groups.

Instrumentation

Alabama's *High School Basic Skills Exit Exam*, Second Edition, and the PACE Competency Cabinets were the two instruments used in this study. Alabama's *High School Basic Skills Exit Exam* is a multiple-choice criterion-referenced test in language, reading and math (Fair Test, 1998). Criterion-referenced assessments (CRAs) compare student performance to clearly defined curricular objectives, standards or skill levels and provide information on how well a particular course of study has been learned (Fisher, Roach, & Kearns, 1998).

The PACE Learning Systems instrument was chosen for remediation purposes because of its comparatively low cost and, more importantly, its congruency with the competencies presented on the *High School Basic Skills Exit Exam*. PACE Learning Systems provided a cross-reference that delineated the competencies included on the *High School Basic Skills Exit Exam* and identified the PACE modules, from their cabinets, that were designed to remediate those competencies. For the 23 reading competencies, 50 modules were provided. For the 23 language competencies, 25 modules were provided; and for the 25 mathematics competencies, 40 modules were provided. Because of the similarities between the skills necessary to work through the PACE modules and those needed to complete the pencil and paper *Exit Exam*, the Pace Competency Cabinets were deemed an appropriate choice for tutorial and remediation efforts.

Sample

Participants in this study were 270 junior and senior students at Decatur High School, Decatur, Alabama, who failed at least one portion of Alabama's *High School Basic Skills Exit Exam* from 1994-1999. Although 366 students qualified to participate, the sample consisted of 270 students. Ninety six students were removed from consideration because only a beginning test score was available. All 270 students were offered the opportunity to participate in remediation classes.

Of the 270 students who were offered remediation, 134 declined. These students constituted the comparison group. Some students failed more than one section of the *Exit Exam*, therefore the comparison group sample size varies across discipline. For reading the comparison group consisted of 56 students, 72 students

comprised the math comparison group, and 112 students served as the comparison group in language.

Students who volunteered, with parental consent, to participate in at least one portion of the remediation classes comprised the treatment group ($n = 136$). Again, because some students failed more than one section of the exam the sample size varies across discipline. The treatment group for reading consisted of 33 students, 52 students comprised the math treatment group, and 99 students served as the language treatment group.

A comparison of demographic data of the total school population to the comparison and treatment groups reveals several points of interest. Although 51% of Decatur High School's student population was male, 59% of the students who failed the *Exit Exam* were male (See Table 1). As suggested in the literature, those students considered most at risk of failing were minorities and those of low socioeconomic status.

Data Collection

The data were collected from Decatur High School student archives for students who failed any portion of Alabama's *High School Basic Skills Exit Exam* from the fall testing in 1994 through the spring testing in 1999. Each subject was assigned a number to ensure privacy. The information collected consisted of gender, race/ethnicity, and socioeconomic status as determined through records of participation in the federally funded free or reduced lunch program (Holman, 1995; Ivory, 1993), and all previous dates and scores on the *High School Basic Skills Exit Exam*.

In order to accurately analyze the data, a beginning test score (pre-test) and an ending test score (post-test) was required from each of the 270 subjects. For the comparison group, the scores from the first attempt on the *Exit Exam* was recorded as the pre-test score, and the scores from the final attempt was designated as the Post-test score for data collection. For the treatment group, the attempt on the *Exit Exam* prior to enrolling in the remediation class was specified as the pre-test score. The final attempt on the *Exit Exam* was considered the post-test score.

Treatment

Students who failed one portion of the *High School Basic Skills Exit Exam* were given the opportunity to participate in a tutorial remediation program that used the PACE Competency Cabinets. Classes were established in reading, language, and math. Each class followed the format provided by the PACE Learning Systems. A PACE pretest was administered to each student in the appropriate subject area. The student and the teacher collaboratively developed an individual study schedule based on the results of the pretest and specific areas of deficiency on the *High School Basic Skills Exit Exam*.

Remediation materials consisted of modules, which correlated with the competencies tested by the *High School Basic Skills Exit Exam*. Upon completion of each pencil and paper module, practice tests were administered. A score of 90 indicated the student's readiness to take a module test. PACE provided two tests of equal difficulty for each module. A score of 90 on the module test indicated mastery

of that competency; however, with a score less than 90, the student repeated the module and took the alternate test. A module could be repeated as many times as necessary for mastery.

PACE Competency Cabinets included computer software that contained materials identical to the pencil and paper modules. Only after completing the pencil and paper modules could students elect to work on computers to enhance their learning.

Methodology and Data Analysis

An ex post facto research design was used to examine the reading, mathematics, and language scores of the treatment and comparison group. Prior to analyzing posttest scores, pretest scores were assessed for equivalence using a 2x2x2 factorial analysis of variance (ANOVA). Because gender and ethnicity differences in posttest scores were of interest, these two variables were included in the analyses along with group membership (participation or nonparticipation in PACE). For the ethnicity variable, only African American and white students were compared. Separate analyses were conducted for the three subject areas of the Exit Exams (reading, mathematics, and language). In order to determine equivalence between the two groups, the two-way interaction effects between group membership and ethnicity and group membership and gender were evaluated at $\alpha = .10$. The main effect of group membership was evaluated at $\alpha = .05$. ANOVA results are presented in Table 2.

The difference in the mean pretest reading score for the treatment group ($M = 59.7$, $SD = 5.5$) and the comparison group ($M = 57.3$, $SD = 8.5$) was not statistically significant and no interaction effects were found (see table 2). Therefore, equivalence was established for the reading analyses. However, an interaction effect was found between group membership and race ($F(1, 119) = 7.38$, $p = .008$) for the pretest mathematics scores and a main effect of group membership was found for the pretest language scores ($F(1, 186) = 7.5$, $p = .01$).

Because of the presence of an interaction effect for group membership and race, interpretation of posttest results would be problematic. Therefore, the decision was made to exclude students scoring below 35 on the pretest (the lowest score for students in the treatment group). Re-analysis of the data showed no group by ethnicity interaction effect ($F(1, 103) = 1.87$, $p = .175$) and no main effect for group membership ($F(1, 103) = 1.55$, $p = .216$), thus establishing a more appropriate comparison group for posttest mathematics results. For analyses of posttest scores, the sample size for the mathematics comparison group was decreased from 72 to 59.

Similarly, in order to establish more comparable groups for interpreting posttest results in language, only students scoring above 41 on the pretest (the lowest score for the treatment group) were included in analyses. When the lowest scoring students in the comparison group were excluded, differences in the pretest language means were found to be statistically non-significant ($F(1, 183) = 2.14$, $p = .145$). Therefore, for analyses on posttest scores, the language comparison group sample size was decreased from 105 to 92.

For posttest analyses, chi-square statistics were assessed to determine if significant differences existed between the proportion of students passing the Exit

Exam for the treatment and comparison groups. Gender and ethnicity differences were assessed using chi-square statistics for each group. Results were evaluated at $\alpha = .05$. Estimates of effect size using the arcsine transformation procedure (Cohen, 1977) are also reported. Frequency counts for each subgroup are presented in table 3.

Results

Reading

Of the 33 students who participated in the PACE reading training, 25 (75.8%) passed the reading portion of the *Exit Exam*. This proportion is statistically significantly higher than the 46.4 % success rate of the comparison group ($\chi^2(1, N = 89) = 7.3, p = .007$). The estimated effect size is .60.

No difference was found between the two groups for the passing rates of female students ($\chi^2(1, N = 48) = 2.35, p = .125$). However, the proportion of male students passing the reading exam was statistically significantly higher for the PACE participants ($\chi^2(1, N = 41) = 6.72, p = .01$). The passing rate for white PACE participants ($n=8$) was 75% and for white students in the comparison group ($n=26$) the passing rate was 50%. The small sample size for the PACE participants makes interpretation of the chi-square test problematic; however, Fisher's exact test indicates no statistical significance ($p = .62$). For African American students, the proportion of PACE participants who passed was 76% and the proportion passing in the comparison group was 43.3% ($\chi^2(N = 55) = 5.98, p = .014$).

Mathematics

After excluding students who scored lower than 35 on the Mathematics pretest, the treatment group and the comparison group had a sample size of $n_t=52$ and $n_c=59$, respectively. Of the students who participated in the PACE mathematics training 78.8% passed the mathematics section of the *Exit Exam*. In contrast, the passing rate for the comparison group was 57.6 % ($\chi^2(1, N = 111) = 5.68, p = .017$). The estimated effect size is .5.

No difference was found in the proportion of male students passing between the two groups ($\chi^2(1, N = 67) = .485, p = .486$). For female students, the passing rate for PACE participants was 89.3% and for non-participants the passing rate was 57.7% ($\chi^2(1, N = 54) = 7.01, p = .008$). There was no difference between the two groups in the proportion of white students passing the mathematics exam ($\chi^2(1, N = 53) = 3.59, p = .058$). The proportion of African American students passing was 74.4 % for participants and 45.5 % for the comparison group ($\chi^2(1, N = 58) = 5.09, p = .024$).

Language

Excluding students whose score on the Language pretest was lower than 41 left a treatment group samples size of $n_t=99$ and a comparison group sample size of $n_c=92$. The treatment group passing rate on the language section of the *Exit Exam* (82.5%) was statistically significantly higher than the 46.8% passing rate for the

comparison group ($\chi^2(1, N = 191) = 26.67, p < .001$). The estimated effect size is .8.

For male and female students, significant differences were found in the passing rates of the two groups. The passing rate for male PACE participants was 77.6 % compared to the 43.9 % passing rate of male non-participants ($\chi^2(1, N = 123) = 14.5, p < .001$). For female PACE participants, 89.7 % passed the language exam whereas 48.3% of the female non-participants passed ($\chi^2(1, N = 68) = 15.5, p < .001$). Significant differences were also found for African American and white students between the two groups. The passing rate for African American PACE participants was 76.9%; for non-participants, 31.6 % of the African American students passed the language *Exit Exam* ($\chi^2(1, N = 100) = 20.58, p < .001$). Of the 32 white students in the treatment group, 93.8% passed the exam while only 4.2% of the white students in the comparison group passed ($\chi^2(1, N = 91) = 15.5, p = .001$).

Summary of Results

For each section of the *Exit Exam*, the passing rate for the students who participated in the remediation class was statistically significantly higher than for the students who did not participate. These results are consistent with the findings of Abbott, et. al. (1982). Interestingly, male students seemed to benefit more from the reading remediation whereas female students seemed to benefit more from the mathematics remediation. The passing rate for African American students in the treatment group was higher than for African American students in the comparison group on all three exam components.

Examination of the archival data revealed that the students who volunteered for remediation had pretest scores that were higher, on average, than students who chose to decline help, particularly for mathematics and language. This suggests that other factors (for example, motivation to succeed) are important considerations. As a preliminary assessment of the PACE cabinets for remediation, the positive findings are promising.

Implications

In 1991, the U. S. Secretary of Labor Lynn Martin convened a panel to examine the role of schools in preparing students for work. *The Secretary's Commission on Achieving Necessary Skills (SCANS)* reported that schools continue to train students for a work place that no longer exists in American society. In order to reestablish its place as the premier educator of children, public schools had to raise standards to meet the challenge; and students had to work harder to graduate (US Department of Labor, 1991).

For several years, students in many other states have been required to pass competency exams as a prerequisite for promotion and/or graduation. There is a renewed pressure from the public to improve education, including showing student competence on standardized tests. *NCLB* now has become the centerpiece of the federal approach to improving education. It requires that states establish statewide proficiency and schools to demonstrate student progress toward minimum levels of

proficiency tied to sanctions for lack of progress and rewards for success (Department of Education, 2002).

Proponents of high-stakes testing have identified several advantages of the tests. They claim that poor and minority students stand to benefit most from a strong accountability system that includes annual testing. Though there is considerable controversy, most believe high stakes tests to be objective, impartial, and lessen the potential for conflict between a teacher and a student (Kucirik, 2002). They provide a basis for comparing students nationwide. Teachers gain a sense of focus and purpose for their lessons, and students are given tangible incentives to perform well. The public approves of standardized tests and regards the results as valid (Madaus, 1991).

With such support for standardized testing across the nation it is imperative that educators get creative and innovative in designing means by which all students have the opportunity to succeed on these standardized tests. The new federal mandate in the form of *NCLB* further reinforces the need for bold actions by educators.

Since there is very little research addressing the effectiveness of remediation as an intervention strategy, educators are left with few, if any, research based options on how to successfully serve students who do not succeed on their first attempt to pass these tests. In this setting, there was some success for students who participated in the PACE Competency Cabinets remediation program. One must remember that the Pace Learning Competency Cabinets are closely aligned to the Alabama *High School Basic Skills Exam*. However, since students will continue to be required to pass these kinds of standardized tests for promotion and graduation from high school. In many states research on the degree to which Pace Learning instructional materials address a particular set of standards is judicious since the Pace Learning Competency Cabinets have been recently introduced in Georgia.

No group should find themselves cut out of the bounty the United States offers its citizens because of the inability to successfully navigate a single test. Until there is a more rational approach to accountability, schools must take responsibility for creating diverse opportunities so that students at greatest risk have a realistic opportunity for success. It is incumbent on teachers and administrators to develop programs that enable students to bridge the gap.

Further Research

Most of today's educational reforms focus on testing and higher academic standards which led to today's controversial accountability movement. Surprisingly, there has been very little study of the consequences of the movement (Dee, 2002). The American Educational Research Association has recently issued guidelines to educational leaders and educational policy makers about the issue of high-stakes tests. These guidelines suggest there are many opportunities for researchers to look at the consequences of this reform effort as suggested by Dee.

Further research into alternative approaches to remediation is necessary because students in Alabama, Georgia, and other states will continue to be required to pass these kinds of standardized tests for promotion and graduation from high school. Without these alternatives even more students will be left out of educational and economic opportunities afforded by a high school diploma.

References

- Abbott, G. A., Connell, P., Davidson, J., & McLean, J. E. (1982, November). *Minimum competency testing: Inception through remediation*. Paper presented at the annual meeting of the Alabama Personnel and Guidance Association, Mobile, AL. (ERIC Document Reproduction Service No. ED 226 022)
- Airasian, P. W. (1988). Symbolic validation: The case of state-mandated, high-stakes testing. *Educational Evaluation and Policy Analysis, 10*, 301-313.
- Airasian, P. W. (1993, April). Policy-driven assessment of assessment-driven policy? *Measurement and Evaluation in Counseling and Development, 26*, 22-29. (ERIC Document Reproduction Service No. EJ 465 062)
- Anders, P., & Richardson, V. (1992). Teacher as game-show host, bookkeeper, or judge? Challenges, contradictions, and consequences of accountability. *Teachers College Record, 94*, 382-395.
- Archer, E. L., & Dresden, J. H. (1986). *A new kind of dropout: The effect of MCTing on high school graduation in Texas*. San Francisco, CA: American Educational Research Association. (ERIC Document Reproduction Service No. ED 269 454)
- Boone, W. J., Rogg, S. R., Kahle, J. B., & Damnjanovic, A. (1997). *Race, gender, test length, and missing data. Why estimates of performance may be clouded*. Paper presented at the National Association for Research in Science Teaching Conference, Chicago, IL. (ERIC Document Reproduction Service No. ED 406 187)
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.
- Dee, Thomas. *Education Next*. Retrieved May 3, 2004, from http://www.findarticles.com/cf_dls/m0MJG/3_3/104835537/pl/article.jhtml.
- Education World. (2000). *School issues: Cautions about high-stakes tests*. Retrieved May 2, 2004, from www.education-world.com/a_issues/issues110.shtml.
- FairTest. (1998). *Testing our children: Alabama*. Retrieved (January 12, 2000), from <http://www.fairtest.org>.
- Fisher, D., Roach, V., & Kearns, J. (1998). *Statewide assessment systems: Who's in and who's out?* Retrieved December 6, 1999, from <http://www.ecs.org>.
- Griffin, B. W., & Heidorn, M. H. (1996). An examination of the relationship between

MCT performance and dropping out of high school. *Educational Evaluation and Policy Analysis* 18, 243-252.

Hampton, E. F. (1997). *Educators' perceptions of the effects of state-mandated testing in a Texas school district and a New Mexico school district along the Mexican border*. Unpublished doctoral dissertation, New Mexico State University, Las Cruces.

Ingram, T.E. (1995). *The Handbook of Administrative Procedures for the High School Basic Skills Exit Exam*. Montgomery, AL: Alabama Department of Education.

Ivory, G. (1993). *Investigating ethnic bias in the mathematics portion of the exit-level Texas assessment of academic skills*. El Paso, TX: Ysleta Independent School District. (ERIC Document Reproduction Service No. ED 362 574)

Kohn, A. (2002). High stakes testing as educational ethnic cleansing. *Education Digest*, 66, 13-14.

Kucerik, E. (2002). The No Child Left Behind Act of 2001: Will it live up to its promises? *Georgetown Journal on Poverty Law and Policy*, 9, 479-487.

Lam, T.C. (1995). *Fairness in performance assessment*. (ERIC Document Reproduction Service No. ED 391 982).

Lawton, M. (1997). *Discrimination claimed in Texas exit-exam lawsuit*. Retrieved January 23, 2000, from <http://www.edweek.org>.

Lindsay, D. (1994). *No racial bias found in Ohio's school exit test*. Retrieved January 23, 2000, from <http://www.edweek.org>.

Madaus, G. F. (1991). The effects of important tests on students. *Phi Delta Kappan*, 73, 226-231.

Manzo, K. K. (1997). *High-stakes: Test truths or consequences*. Retrieved January 23, 2000, from <http://www.edweek.org>.

National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: US Department of Education.

Samuelsen, S., (2001). Student testing: The stakes are rising. *State Legislatures*, 27, 30-34.

State Department of Education. (1996). *Laws of Alabama relating to education*. New York: Lawyers Cooperative Publishing..

Tuckman, B. W. (1994). *Conducting educational research* (4th ed.). Fort Worth, TX: Harcourt Brace.

U. S. Department of Education. (1996). *What the research says about student assessment*. Retrieved January 26, 2000, from <http://www.ed.gov/pubs>.

U. S. Department of Education. (2002). *The no child left behind act of 2001: Executive summary 1*. Retrieved May 3, 2004, from <http://www.ed.gov/offices/OESE/esea/summary/html>.

U. S. Department of Labor. (1991) What work requires of schools: A SCANS report for America 2000. *The Secretary's Commission on Achieving Necessary Skills*, vii.

Vanneman, A. (1998a). Long-term trends in student mathematics performance. *NAEPfacts*, 3(2).

Vanneman, A. (1998b). Long-term trends in student science performance. *NAEPfacts*, 3(3).

Vanneman, A. (1998c). Long-term trends in student writing performance. *NAEPfacts*, 3(4).

APPENDIX

Table 1

Decatur High School, Decatur, Alabama, Mean Student Demographics from 1994-1999

	Population	Failing <i>Exit Exams</i>	Comparison	Participant Means
Average Enrollment	1088	280	141	139
Male	51%	59%	63%	56%
Female	49%	41%	37%	44%
White	70%	48%	58%	34%
Male	36%	29%	37%	19%
Female	34%	19%	21%	15%
Black	26%	52%	39%	63%
Male	13%	52%	24%	34%
Female	13%	23%	15%	29%
Other	3%	N/A	3%	3%
Male	2%	N/A	2%	3%
Female	1%	N/A	1%	0%

Table 2

ANOVA Results for Pretest Scores

<i>Reading</i>				
Source	df	F	η	p
Group	1	3.4	.04	.07
Gender	1	.01	.001	.91
Race	1	1.7	.02	.19
Group by gender	1	1.2	.02	.28
Group by race	1	.38	.01	.54
Group by Gender by Race	2	.003	.003	.96
S within group error	81	(58.1)		
<i>Mathematics</i>				
Group	1	6.9	.06	.01
Gender	1	.26	.002	.61
Race	1	6.2	.05	.02
Group by gender	1	.19	.001	.66
Group by race	1	7.4	.06	.008
Group by Gender by Race	2	.24	.003	.63
S within group error	119	(101.2)		

Language

Source	df	F	η	p
Group	1	7.6	.03	.01
Gender	1	2.5	.01	.11
Race	1	4.9	.04	.03
Group by gender	1	.09	.001	.76
Group by race	1	.24	.01	.63
Group by Gender by Race	2	.09	.001	.91
S within-group error	196	(56.9)		

Note. Values enclosed in parentheses represent mean square errors. S = subjects.

Table 3

Comparison of the number of students passing the *Exit Exams*.

	Treatment		Comparison	
<i>Reading</i>	Pass	Fail	Pass	Fail
African American	19	6	13	17
White	6	2	13	13
Female	14	7	12	15
Male	11	14	1	15
Total	25	8	41	48
<i>Math</i>				
African American	29	10	10	12
White	12	1	24	13
Female	25	3	15	11
Male	16	8	19	14
Total	41	11	34	25
<i>Language</i>				
African American	50	15	11	24
White	30	2	32	27
Female	35	4	14	15
Male	45	13	29	36
Total	82	17	43	49