DOCUMENT RESUME

ED 460 155 TM 033 629

AUTHOR Mulvenon, Sean W.; Connors, Joanie V.; Lenares, Denise
TITLE Impact of Accountability and School Testing on Students: Is There Evidence of Anxiety?
PUB DATE 2001-11-00
NOTE 32p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (30th, Little Rock, AR, November 13-16, 2001).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Accountability; *Anxiety; *Elementary School Students; Intermediate Grades; Standardized Tests; *Student Attitudes; *Testing Programs
IDENTIFIERS Arkansas

ABSTRACT

The marked increases in standardized achievement testing have raised concerns of the educational community regarding the value of standardized testing programs and their potentially harmful effect on students. This study incorporated student perceptions and their views of standardized testing and combined this information with their performance on the exam. Student results from the fall 2000 Stanford Achievement Test, edition 9 and the spring 2000 Fourth Grade Benchmark examination, a criterion-referenced test used in Arkansas, were combined with student survey data about attitudes and perceptions for 283 students. The results suggest that the vast majority of students do not exhibit stress and have positive attitudes toward standardized testing programs. The study also showed no evidence that anxiety or pressure negatively impacts student achievement. (Contains 1 figure, 6 tables, and 33 references.) (Author/SLD)

Running head: Test Anxiety

Impact of Accountability and School Testing on Students: Is There Evidence of Anxiety?

By

Sean W. Mulvenon, Ph.D.,

Joanie V. Connors, Ph.D.,

and

Denise Lenares, M.Ed.

University of Arkansas

Address Correspondence to: Sean W. Mulvenon
254 Graduate Education Building
College of Education and Health Professions
University of Arkansas
Fayetteville, AR 72701
Phone: (501) 575-8727
Email: seanm@uark.edu

Abstract

The marked increases in standardized achievement testing have raised concerns of the educational community regarding the value of standardized testing programs and their potentially harmful effect on students. This study incorporated student perceptions and their views of standardized testing and combined this information with their performance on the exam. The results suggest that the vast majority of students do not exhibit stress and have positive attitudes towards standardized testing programs. Also, there is no evidence found in this study that anxiety or pressure negatively impacts student achievement.

Impact of Accountability and School Testing on Children

Introduction

A major policy change in education, originated in the 1970s and resulted from growing political pressure on public schools and teachers to be responsible, or more accountable, for the academic achievement of their students (Fairchild & Zins, 1986). This movement, often referred to as school accountability, gained additional momentum with the anemic performance of U.S. students on several large national and international studies, such as the Third International Mathematics and Science Study (Kimmelman,, Kroeze, Schmidt, van der Ploeg, McNeely & Tan, 1999). Political pressure to address what was perceived by many as a struggling educational system, resulted in a new generation of standardized tests for the purpose of evaluating school systems, from district to classroom levels to "demonstrate" students are learning (Bernauer & Cress,1997). The use of these test results to make important decisions about teachers' jobs and pay, school funding, and the promotion of students to higher grades has become known as "high stakes testing" (Kubiszyn & Borich, 2000).

Numerous articles have suggested that there is a damaging effect on students associated with the influx of standardized testing and implementation of high-stakes accountability programs (Bernauer & Cress, 1997; Gordon, 2000; Paris, 1992; Shepard, 1991; Wolk, 2001). However, these articles have been based on opinions and anecdotal evidence at worst and at best they have been based on unidimensional surveys that evaluate the impact of testing without linking results to student performance on standardized tests.. The purpose of this study was to formally evaluate the impact of anxiety and pressure on student performance by combining student perceptions and attitudes about standardized tests with actual performance on these exams.

Among this growing number of articles about the impact of testing on student achievement, Baker, O'Neil and Linn (1996) reported that less than 5% of articles about high-stakes testing program are research based. Thus, much of the literature appears to consist of papers discussing the theoretical advantages or disadvantages of high stakes testing rather than formal investigations of the impact accountability programs. The opinions can be separated into two views, those reporting a negative impact of testing on teachers, students and schools (Bernauer & Cress, 1997; Gordon, 2000; Paris, 1992; Shepard, 1991) and those who advocate on behalf of testing programs to bring needed reforms to the U.S. educational system (American Federation of Teachers [AFT], 2000; Cizek, 2000; Zirkel, 1999).

Articles which suggest a negative impact of the expanded testing and accountability programs say that teachers will 'teach to the test' and neglect other curriculum and higher order thinking skills that are less easily measured in standardized tests (Jones, Jones, & Hardin, 1999; Mooney, 1996; Paris, 1992; Shepard, 1991). Other concerns cited include lost instructional time in the classroom (due to testing and test practice sessions), the negative impact on students who perform poorly on these exams, yet have demonstrated higher levels of achievement via more subjective measures of performance (e.g., written work and class participation.), and the tests not measuring the effectiveness of the prescribed curriculum (Etsey, 1997; Paris; Shepard).

Advocates of expanded testing and accountability programs typically express a concern over the lack of a "standardized" assessment instrument that allows for direct comparisons of educational programs (Linn, 2000). Further, several studies (Bejar & Blew, 1981; Cizek, 2000; Kuh & Hu, 1999; Landrum, 1999; Ziomek & Svec, 1995) have suggested a pervasive problem with grade inflation in our

school systems. These studies suggest little value can be placed on grade point averages as the universal indicator of achievement with such inherent variability and subjectivity in the assignment of grades by instructors. Standardized tests help to overcome this issue by providing a common metric for assessing student achievement when these students come from a variety of educational and situational environments.

Effects of Testing on Children

Many are concerned that testing will ultimately be harmful to children with testing practices and test scores becoming more integral, influential, and important in how their lives progress (Paris, 1992). As a result, several studies have been conducted measuring the influences of testing on a child's academic achievement and their attitudes towards testing (Karmos & Karmos, 1984; Paris, Lawton, Turner, and Roth 1991; Paulson, Marchant, & Rothlisberg, 1998; Wright, 1999).

Paris, Lawton, Turner, and Roth (1991) surveyed students from Grades 2 - 11 for their attitudes about testing. Their results suggest three general trends; growing disillusionment about tests, increased use of inappropriate test-taking strategies, and decreased motivation and effort when taking tests. "A large number of students, especially low achievers, become anxious about tests, cheat, try halfheartedly, or use poor test-taking strategies" (p. 12). The Paris et al., study is representative of many of the efforts to evaluate the impact of standardized testing, with the conclusions suggesting an anxiety, malfeasance on the part students, or test taking strategies as the result of standardized testing programs.

Socioeconomic Status (SES) also has been suggested as a factor in student achievement. Wright (1999) surveyed 3rd and 4th graders in 33 elementary schools and found the risk factors of

ethnic minority status and low family income" (p. 350) had the greatest effects on achievement. Sweetland and Hoy (2000) found that SES and teacher empowerment significantly contributed to the reading and math achievement of students. Caldas and Bankston (2000) found that SES was less important than number of single parent families in a school for impacting achievement scores of students.

At least two studies have asserted that testing has negative effects on children without noting that their data source was teachers' opinions about what children felt about testing and did not ask the children themselves (Jones, Jones & Hardin, 1999; Mehrens, 1998). Information obtained from students' perspectives on testing appears to have been neglected in general (Paris, 1992).It is important to get the perspectives of the students themselves instead of through the eyes of school personnel because the nature and meaning of testing can be very different for these two (Weinstein & Middlestadt, 1979).

Many of these studies based on teacher opinions reported testing had negative effects on children, but neglected to examine biases coming from the teachers' perspectives. Smith and Rottenberg (1991), concluded from teacher reports that testing caused children anxiety and challenged the self-concept of the lower skilled students (cited in Mehrens, 1998). It is likely that this evidence for negative effects on children that comes from teacher reports is shaped by teachers' needs and personal reactions regarding testing programs (Shepard, 1991). Further, numerous studies specifically evaluating teacher attitudes about standardized testing consistently report strong dissension on the part of teachers toward these programs (Urdan & Paris, 1994; Bliem & Davinroy, 1997). Thus, asking teachers to evaluate the feelings, mental state, or pressures experienced by students may produced extremely

biased interpretations driven by their own personal beliefs about testing.

Test Anxiety

Some studies suggest older students (secondary level) suffer from anxiety about standardized testing and this is especially likely if the student is low achieving (Karmos & Karmos, 1984; Paris, 1992; Paris et al., 1991). According to Paris (1992), older students experience increases in anxiety in reaction to evaluations when they are unsure of their own ability, and that in-turn detracts from their motivation. Further, Paris suggests the inability to perform well on a test contributes to increased anxiety toward other evaluations. Paris et al., also found a relationship between low achievement and test anxiety in students. These studies have reported an association between low test scores and test anxiety, however, all fail to address if this issue is more associated with the student is a low performing student or if low tests scores are purely attributable to test anxiety.

In their study of the relationship between students' achievement test performance and their attitudes towards standardized achievement tests, Karmos and Karmos (1984) found student attitudes about testing were moderately positive. However, of the 350 6th to 9th graders surveyed, 25% responded negatively to 5 of the 12 items about the use of test results and these attitudes were significantly related to achievement levels. They concluded there was a relationship between negative attitudes towards standardized tests and test anxiety for this subgroup.

School Climate

School climate has been referred to as "... psychological factors in the school context affecting student learning and general well-being" (Esposito, 1999, p. 366). School climate has been described as having a mediating effect on achievement scores because it affects the way student take in and

understand knowledge (Paulson, Marchant & Rothlisberg, 1998). More specifically, the situational context of the testing environment may contribute, either positively or negatively, to student achievement.

Sweetland and Hoy (2000) evaluated instruments to assess school climate and found two dimensions being measured; openness and healthiness that discriminated positive school climate from negative. While these dimensions were not equal, they were found to be related "Open schools tend to be healthy and heathy schools tend to be open." (p. 708). Sweetland and Hoy's own study found four factors of school climate that explained 71% of the variance in climate; collegial relationships among teachers, collegial relationships between teachers and administrators, academic press (pressure to learn), and community interest.

The issue of climate is becoming more relevant with teachers reporting high-stakes testing and accountability programs creating greater pressure to raise test scores and a resultant negative work climate (Barksdale-Ladd & Thomas, 2000). In an educational environment, teachers exhibiting this pressure to students, either in an overt or covert fashion, may adversely impact student achievement. Basically, teachers may inadvertently place additional pressure on students to perform well on the exam. This may be occurring due the belief by many teachers that poor test scores will translate to lower wage increases or termination a perceived negative climate based on the desire to increase test scores (Kubiszyn & Borich, 2000).

Pressure

School climate that was characterized by strong academic press (pressure to learn) was found to account for 50% of the variance in achievement scores in a study by Goddard, Sweetland and Hoy

(2000). Phillips (1997) found a climate of academic press predicted mathematics achievement as well as attendance.

Several studies have been conducted looking at the significance of the influence of teacher expectations and behavior on students' academic performances in standardized and classroom assessments. Brookhart and DeVoge, (1999), when investigating the role that classroom assessment plays in student motivation and achievement, found the way in which teachers communicate their performance expectations for students helps students to form an idea of what is important to learn.

Teacher behavior can also influence students' self-perceptions of their abilities. Weinstein and Middlestadt (1979) investigated high and low achieving male students' perceptions of teacher interactions. The results of the study suggested, "...teacher behavior toward individual students can be seen as providing information about achievement status to the student as well as to peers" (p. 430).

In studying educational achievement testing from the perspective of parents, teachers, political, and students, Paris (1992) has found that most parents have positive attitudes about school testing because it provides evidence of their child's accomplishments, however, many times they do not know how to interpret the scores because they do not receive enough information. Furthermore, some parents worry about how testing will affect their child's learning and motivation especially if their child has experienced test anxiety or failure in the past (Paris). Pressure may also be placed on the student by parents, with the inference the test is more representative of how the students are performing relative to other students.

Rewards

Another concept that may be related to student performance and achievement is the praise and rewards that students receive from parents, teachers and others for their work. Paris (1992) finds evidence that students learn their beliefs about themselves from rewards and praise, which act as reinforcement for developing skills. In many school environments testing day is greeted with few courses, special treats, and recesses. Further, some schools administrators have been known to promise "pizza days" or other rewards for meeting performance goals. Ironically, promises of rewards, something intended to motivate students may actually be contributing to the pressure some students are experiencing to perform well.

Self-Efficacy

Perceived self-efficacy is a concept developed by Bandura (1986) and is defined as perceptions and confidence about one's own abilities. Self-efficacy beliefs have been found to influence motivation, interest, plans for the future and ability to withstand disappointments (Bandura, Barbaranelli, Caprara, & Pastorelli, 2001). Self-efficacy is shaped by parental expectations of the child and Bandura believes that this explains the high influence of SES on student achievement because social class can mediate the expectations of children. Jinks and Morgan (1999) developed a scale of academic self-efficacy and found it was related to higher grades in their study.

Attitude towards Testing

Karmos and Karmos (1984) conducted a study that indicated students' perceptions of the importance of standardized tests were related to their test performance. In their study researching the relationship between the attitudes of students in Grades 6 - 9 and their performance on the Stanford Achievement Test (SAT), they found that "...attitudes of the students...accounted for 14% of the

variance in scores on the SAT" (p.66). They went on to state that researchers should not ignore the importance of students' attitudes when taking into consideration variables that influence achievement test performance.

The Current Study

A comprehensive evaluation of the many factors that may influence student performance on standardized assessments was developed. Students who had just completed the mandated testing with the Stanford Achievement Test-9th Edition (SAT-9) were surveyed for their attitudes and reactions to the testing process. An important aspect of this study was the combination of student results from the Fall 2000 SAT-9 test and the Spring 2000 4th Grade Benchmark Exam with student survey data on attitudes and perceptions about standardized testing.

To examine these issues, a 24 item questionnaire was developed with questions addressing all of the pertinent issues found in the literature, including test anxiety, attitude towards testing, school climate, pressure from parents and teachers, and self-efficacy in math and reading. All of these questions were computed with student scores on the SAT-9 and the Benchmark Exam to examine the factors within these questions which might predict performance on the exams.

Additionally, as a method to cross-validate performance on the SAT-9 and evaluate a consistent effect for test anxiety, student scores for the Spring 2000 4th Grade Benchmark Exam were also merged with Fall SAT-9 performance data. The 4th Grade Benchmark Exam is a criterion referenced exam used as part of the Arkansas Department of Education Accountability System. Students who completed the 5th Grade SAT-9 (a norm referenced test) in Fall 2000 had also completed the 4th Grade Benchmark Exam in Spring 2000. The combination of both norm- and

criterion-referenced information was designed to address teacher concerns for focusing on one test versus another.

The broad purpose of this study was to investigate what impacts student performance on standardized exams other than innate ability. Several research hypotheses were investigated to accomplish this purpose. The first hypothesis was to determine if associations between anxiety, school climate, pressure and standardized test performance are evident. The second hypothesis examined whether student attitude about testing contributes to test performance. The third hypothesis examined the issue of self-efficacy of students for math and reading, and associations with test performance, anxiety, and pressure. The fourth hypothesis was to see if the schools themselves had an impact.

Methods

Subjects

All Students in fifth grade from a southern school district with 10 elementary schools were selected for participation in this study. Students were given a packet that contained an explanation of the study for parents, a parent consent form for participation, and a parent survey. A total of 283 students (over 50%) returned completed permissions forms signed by their parents and were allowed to participate in the study.

Instruments

As part of the normal standardized testing required by the Arkansas Department of Education (ADE), students were administered the Stanford Achievement Test, Version-9 (SAT-9). Additionally, during the preceding spring, as fourth graders, all students had completed the Fourth Grade Benchmark exam, a criterion referenced exam also required in Arkansas. Following the SAT-9 administration, the

283 students with permission forms were administered a survey consisting of 24 questions to obtain information on their views or perceptions about standardized testing.

A factor analysis of the student survey data revealed a six factor solution. The question loadings on the six factor solution were judged to indicate the following constructs: (a) Test Anxiety, (b) Rewards (c) Pressure, (d) Reading Self-Efficacy, (e) Math Self-Efficacy and (f) Attitude Towards Testing. Scales were constructed from these six factors. The internal consistencies, computed using Cronbach's Alpha, and intercorrelations among five of the constructs are reported in Table 1. The Cronbach alpha values range from .65 to .85, and thus meet or exceed the .65 threshold typically expected for latent constructs. The sixth factor, Attitude Towards Testing, produced an internal consistency value of .54, thus this factor was excluded from further analysis in this study. The intercorrelations reveal an interesting result, with only one relationship, Pressure and Test Anxiety with a value of .55, having a Pearson correlation coefficient that exceeds .20. This indicates that approximately 30% of the variance in Test Anxiety can be associated with Pressure [$R^2 = (.55)^2$].

Procedures

The selected school district administers the SAT-9 during a one-week period in September, in adherence to ADE testing guidelines. Elementary school principals were approached, and approval to obtain information from students during exam week was granted from nine of ten principals. The lone dissent was from a principal who indicated there was too much distraction during the week of testing to subject students to any additional surveys. The remaining principals allowed administration of the student survey to all eligible participants immediately after the students had completed testing, on a day of their choosing, during test week.

Table 2 contains the information on the correlations among the standardized test scores and the associations with the latent constructs. Pressure has the strongest relationship with the SAT-9 reading score, with a correlation of .29. It is interesting to note, even at this preliminary point in the study, the consistent trends of direction and strength between standardized test performance and the survey constructs.

Results

The purpose of this study was to investigate if latent constructs, such as test anxiety, school climate, pressure, rewards, attitudes towards testing, reading self-efficacy, and math self-efficacy impact student performance on standardized exams. A series of analyses to investigate specific secondary hypotheses were completed to provide information on this research question.

Hypothesis 1: The levels of anxiety, school climate, rewards and pressure are associated with standardized test performance.

A series of multiple regression analyses including the CRT Literacy, CRT Mathematics, SAT-9 Reading, and SAT-9 Mathematics scores on Anxiety, School Climate, Rewards and Pressure were completed. Using a nominal alpha level of .10 and a per analysis alpha of .02, produced $F(6,201) = 2.74$ ($p < .0138$), $F(6,201) = 3.55$ ($p < .0023$), $F(6,201) = 4.60$ ($p < .0002$), and $F(6,201) = 4.50$ ($p < .0003$) for CRT Literacy, CRT Mathematics, SAT-9 Reading, and SAT-9 Mathematics scores, respectively.

Using an alpha of .02, review of the four predictors for statistical significance identified Pressure with $t = -2.04$ ($p < .0425$) for predicting reading performance on the SAT-9 exams. The test of significance revealed $t = -2.00$ ($p < .0471$) for Pressure predicting performance on the mathematics

scale of the SAT-9. Pressure with $t = -2.37$ ($p < .0188$) was also an important predictor of Literacy performance on the Benchmark Exam. This process, repeated for Benchmark Exam mathematics performance identified Rewards with $t = -2.08$ ($p < .0385$) as predicting student outcomes. The remaining predictor variables did not contribute an important amount of variance explained in the model.

The unique amount of variance explained by each of the preceding "statistically" significant results is less than 2%. Thus, even though these variables are statistically significant, the practical interpretation of these results is they do not have a great impact on student achievement, with little variability in achievement scores actually predicted with climate and rewards. Regardless, given the relationship between the predictors and test scores, students with more favorable attitudes toward testing tended to have higher scores. Further, this group represents about 75% of students participating in the study.

Hypothesis 2: Student attitude towards testing will predict performance on achievement tests.

The Attitude Towards Testing scale, as previously stated did not produce a construct that was internally consistent. Subsequently, any analysis utilizing this construct would produce inconclusive outcomes. However, it appears that the lack of results for this scale may be due to the complexity of student attitude towards testing and not to a lack of relationship between attitudes and achievement. This became apparent when responses to questions 15, 16, and 20 were computed as forced dichotomies. Students with scores of three or greater were deemed as "positive" in their attitudes and those with scores less than three deemed "negative."

For example, responses to question 15 (perceived value of the SAT-9) were significantly related to student exam performance when positive and negative responses were compared. Student responses to question 15 produced statistically significant differences on the Benchmark Math of $F(1,202) = 4.92$ ($p < .0276$), SAT-9 Reading with $F(1,228) = 9.70$ ($p < .0021$), and SAT-9 Math with $F(1,227) = 15.63$ ($p < .0001$) (see Table 6). The effect sizes were .36, .48, and .61, for Benchmark Math, SAT-9 Reading and SAT-9 Math tests, respectively.

The results for question 16 (I believe the SAT-9 motives me to learn) did not produce any statistically significant differences or any effect sizes greater than .20. Subsequently, this item failed to produce any identifiable trends in performance based on a motivation to perform well, with students of all ability levels represented in each category of the question.

Question 20 (I like test week) produced statistically significant results, but due to violations in the assumption of homogeneity of variance, they were not reported. The main issue for this assumption not holding was the discrepant scores for those students with "negative" convictions. Basically, there are some very high performing students who do not like test week and this creates a greater variability in this group and subsequently a violation in the assumption when compared to the "positive" group. The effect sizes were computed for each test, Benchmark Literacy and Mathematics, as well as SAT-9 Reading and Math, and range from .21 to .38, indicating moderate relationships are present and a definite need for further study and refinement of this construct..

Hypothesis 3: Self-Efficacy is associated with student performance on standardized exams

Student scores on Math and Reading Self-Efficacy scales were generally predictive of performance in this study. A second series of ANOVA analyses using student Reading Self-Efficacy

items (questions 23 and 24) to examine the impact on Benchmark Literacy and SAT-9 Reading scores. Student Math Self-Efficacy items (questions 21 and 22) were also examined for Benchmark Mathematics and SAT-9 Mathematics achievement scores (see Table 3).

The results for Reading Self-Efficacy were $F(4,215) = 4.51$ ($p < .0016$) and $F(4,248) = 3.47$ ($p < .0089$) for Benchmark Literacy and SAT-9 reading, respectively. The comparisons for Reading Enjoyment or Q24 were not statistically significant. The results for Math Self-Efficacy were $F(4,215) = 6.07$ ($p < .0001$) and $F(4,248) = 7.76$ ($p < .0001$) for Benchmark Mathematics and SAT-9 Math, respectively. As with reading, comparisons using Math Enjoyment produced no statistically significant results.

Hypothesis 4: School Effects

A final planned analysis was to determine if differences in students' perceptions existed for Test Anxiety, Pressure, Rewards, Attitudes Towards Testing, Reading Self-Efficacy, and Math Self-Efficacy by the school attended. For these comparisons only one result was statistically significant, for Rewards, but this was due to one school whose scores were dramatically different from the remaining eight schools. Further, this school was permanently closed beginning with the 2001-2002 academic year and all students were transferred to other elementary schools.

Discussion

The primary goal of this study was to evaluate the impact of anxiety, school climate, pressure, rewards and self-efficacy on student performance on standardized tests. Although numerous studies have suggested students experience a high degree of anxiety and pressure associated with standardized

testing, and suggest this might lower student achievement, the results from this study suggest there is little evidence for this.

This study examined if anxiety, school climate, pressure and rewards predicted student achievement. Although pressure, rewards, and math and reading self-efficacy were statistically significant in their relationship with achievment, the actual amount of variance explained was limited, with less than 2% of the unique variance in achievement scores explained by each of these variables. The overall amount of variance explained, or model R-square was only .08 or 8 percent.

Math and Reading Self-Efficacy are also very interesting when examined by school. Those schools with the lowest test scores appear to have the highest reported self-efficacy scores. Possible explanations for the discrepant high self-efficacy may be the instructional environment of these schools, lack of test preparation or test savvy.

Limitations of Current Study

A limitation of this study may have been the inability to access all students in this school district. However, given the large percentage of student who did participate, over 60%, the results are fairly stable as an indicator of student concerns regarding anxiety, pressure, school climate, and rewards in the elementary system in this district. Caution should also be used in generalizing data from one school district to the student population at large.

A second limitation is potential bias in the characteristics of parents who allowed or did not allow their children to participate in this study. Another study currently being completed combines student data with a parent survey and this may provide some information on potential biases. The data

from that study suggests parents' opinions regarding the impact of standardized testing are consistent with student self-report data (Zozone, Stegman, & Ritter, 2001).

Implications for Further Study

The concept of self-efficacy in mathematics and reading needs to be investigated further as they have not been studied enough to have standardized measurement scales. These further refinement of these concepts may help to evaluate their role in student performance on standardized tests.

The concept of pressure from parents and teachers also needs to be examined further, specifically regarding the roles of teachers and parents in building pressure about standardized testing. Since the literature often reports that teachers have expressed negative attitudes in general about testing, it will be important to examine how their attitudes may contribute to pressure, both directly and indirectly.

Conclusion

The present study attempted to discover whether anxiety, school climate, pressure and self-efficacy were significantly related to student performance on achievement tests. Anxiety, school climate, pressure from teachers and parents, and school rewards were not found to be significantly related to aspects of performance on standardized tests. Numerous studies have suggested achievement tests produce negative effects through anxiety and pressure placed on students. The current study certainly included some students who expressed similar levels of high anxiety and pressure to perform, both from parents and teachers. However, those cases were not reflective of the overall student sentiment. This study found that most students appear to experience little or no negative effects from testing.

References

American Federation of Teachers, National Council on Measurement in Education, & National Education Association (1990). *Standards for Teacher Competence in Educational Assessment of Students*. Web: Buros, pp. 1-6.

Baker, E., O'Neil, H. & Linn, R. (1996). Policy and validity prospects for performance-based assessment, *American Psychologist, 48,* 1210-1218.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., Barbaranelli, C., Caprara, G., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories, *Child Development, 72,* 187-206.

Barksdale-Ladd, M. A., & Thomas, K. F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. Journal of Teacher Education, 51, 384-397.

Bernauer, J. A., & Cress, K. (1997). How school communities can help redefine accountability assessment. *Phi Delta Kappan, 1,* 71-75.

Bliem, C. L., & Davinroy, K. H. (1997). Teachers' beliefs about assessment and instruction in literacy. CSE Technical Report 421, National Center for Research on Education, Standards and Student Testing (CRESST). Los Angeles, California.

Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role classroom assessment in student motivation and achievement. *Applied Measurement in Education, 12*, 409-425.

Caldas, S. & Bankston, C. (2000). Multilevel examination of student, school and district-level effects

on academic achievement, *Journal of Educational Research 93*, 91-100.

Esposito, C. (1999). Learning in urban blight: School climate and its effect on the school performance of urban, minority, low-income children.. *The School Psychology Review, 28*, 365-77.

Etsey, Y. K. (1997, March). *Teachers' and school administrators' perspectives and use of standardized achievement tests: A review of published research.* Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Fairchild, T. N., & Zins, J. E. (1986). Accountability practices of school counselors: A national survey. *Journal of Counseling and Development, 65,* 196-199.

Goddard, R., Sweetland, S. & Hoy, W. (2000). Academic emphasis of urban elementary schools and student achievement in reading and mathematics: A multilevel analysis, *Educational Administration Quarterly, 36*, 683-702.

Gordon, B. (2000). On high stakes testing, *AERA Division G Newsletter*, 1-4.

Kaplan, L. S. & Owings, W. A. (2001). How principals can help teachers with high-stakes testing: One survey's findings with national implications. *NASSP Bulletin, 85(622)*, 15-23.

Karmos, A. H. & Karmos, J. S. (1984). Attitudes toward standardized achievement tests and their relation to achievement test performance. *Measurement and Evaluation in Counseling and Development, 17*, 56-66.

Kimmelman, P, Kroeze, D., Schmidt, W., van der Ploeg, A., McNeely, M. & Tan, A. (1999). *A first look at what we can learn from high performing school districts: An analysis of TIMSS data from the First in the World Consortium*, Alexandra, VA: National Science Foundation.

Kubiszyn, T., & Borich, G. (2000). *Educational testing and measurement: Classroom application and practice* (6th ed.). New York: John Wiley.

Jinks, J. & Morgan, V. (1999). Children's perceived academic self-efficacy: An inventory scale, *The Clearing House, 72,* 224-230.

Jones, M. G., Jones, B., & Hardin, B. (1999). The impact of high-stakes testing on teachers and students in North Carolina, *Phi Delta Kappan, 81*, ISSN 0031-7217, 199-203.

Mehrens, W. (1998). Consequences of assessment: What is the evidence? *Education Policy Analysis Archives*, 6, ISSN 1068-2341, 1-24.

Mooney, T. (1996). Do school tests lower standards? *New Statesman, 127*, [On-line]. Available: http://www.findarticles.com/cf_0/m0FQP/n4408_v127/21262293/print.jhtml

Paris, S. (1992). Four perspectives on educational assessment. International Journal of Disability, *Development and Education, 39*, 95-105.

Paris, S. G., Lawton, T. A., Turner, J. C., & Roth, J. L. (1991). A developmental perspective on standardized achievement testing. *Educational Researcher, 20(5)*, 12-20.

Paulson, S. E., Marchant, G. J., & Rothlisberg, B. A. (1998). Early adolescents' perceptions of patterns of parenting, teaching, and school atmosphere: Implications for achievement. *The Journal of Early Adolescence, 18,* 5-26.

Phillips, M. (1997). What makes schools effective? A comparison of the relationships of communitarian climate and academic climate to mathematics achievement and attendance during middle school, *American Educational Research Journal, 34*, 633-662.

Shepard, L. (1991). Will national tests improve student learning? *Phi Delta Kappan*, 232-238.

Sweetland, S. & Hoy, W. (2000). School characteristics and educational outcomes: Toward an organizational model of student achievement in middle schools, *Educational Administration Quarterly, 36*, 703-729.

Urdan, T. C. & Paris, S. G. (1994). Teachers'' perceptions of standardized achievement tests. Education Policy, 8, 137-156.

Weinstein, R. S., & Middlestadt, S. E. (1979). Student perceptions of teacher interactions with male high and low achievers. *Journal of Educational Psychology, 71*, 421-431.

Wolk, R. (2001). Exam anxiety, *Teacher Magazine, 13,* 4.

Wright, D. (1999). Confounded influence on student achievement, *The Journal of Educational Research, 92*, 347-353.

Zozone, M., Stegman, C. & Ritter, G. (2001). Parents' involvement and perception of standardized testing: How does it affect student achievement? In S. Mulvenon (Chair), *The Impact of state testing on student psychological well-being and achievement.* Symposium conducted at the meeting of the Mid-South Educational Research Association Conference, Little Rock, Arkansas.

Table 1

Cronbach Alpha Values for Student Survey Constructs

Construct	n	Number of Items	Cronbach α	Correlations: Test Anxiety	School Climate	Rewards	Reading Efficacy	Math Efficacy
Test Anxiety	236	11	.85	1.00			SYM	
Pressure	251	3	.70	.55*	1.00			
Rewards	236	2	.78	.15	.18	1.00		
Reading Efficacy	257	2	.68	-.16	-.15	.02	1.00	
Math Efficacy	255	2	.65	-.19	-.19	.05	.16	1.00

* Only value that exceeds .30 in value or 10% of variance explained

Table 2

Cronbach Alpha Values and Correlations for Student Survey Constructs

	Correlations								
	SAT-9 Reading	SAT-9 Math	Benchmark Literacy	Benchmark Math	Test Anxiety	School Climate	Reading Rewards	Reading Efficacy	Math Efficacy
SAT-9:									
Reading	1.00								
Math	.76	1.00				SYM			
Benchmark:									
Literacy	.73	.67	1.00						
Math	.69	.74	.68	1.00					
Anxiety	-.19	-.15	-.11	-.19	1.00				
Climate	-.29	-.19	-.22	-.19	.55*	1.00			
Rewards	-.18	-.14	-.06	-.16	.15	.18	1.00		
Reading Efficacy	.26	.13	.19	.15	-.16	-.15	.02	1.00	
Math Efficacy	.03	.23	.18	.19	-.19	-.19	.05	.16	1.00

* Only value that exceeds .30 in value or 10% of variance explained

Table 3

Self-Efficacy and Enjoyment of Subject Results

Question	Variable	F-value	$p <$
Reading Efficacy (q23)			
	Benchmark Literacy	4.51	.0016
	SAT-9 Reading	3.47	.0089
Reading Enjoyment (q24)			
	Benchmark Literacy	0.67	.6099
	SAT-9 Reading	0.65	.6282
Math Efficacy (q21)			
	Benchmark Math	7.76	.0001
	SAT-9 Math	6.07	.0001
Math Enjoyment (q22)			
	Benchmark Math	0.49	.7440
	SAT-9 Math	0.09	.9845

Table 4

Differences in Student Performance and Attitude Towards Testing Items

Question	Test	F	p	Effect Size (Δ)	Means (N)
Q15*	Benchmark Math	4.92	.0276	.36	190(50) vs 223(154)
	SAT-9 Math	15.63	.0001	.61	633(57) vs 658(172)
	SAT-9 Reading	9.70	.0021	.48	651(57) vs 671(173)
Q20**	Benchmark Literacy	--	--	.21	203(50) vs 208(154)
	Benchmark Math	--	--	.26	197(50) vs 221(154)
	SAT-9 Math	--	--	.38	640(57) vs 656(173)
	SAT-9 Reading	--	--	.28	657(57) vs 669(172)

* Only those tests with statistically significant values were reported. Question 16 had no statistically significant test results.

** All model assumptions where checked, and in places where there were violations only effect sizes were reported.

Table 5

Means of Standardized Test Scores for Math and Reading Efficacy by Level of Response

Level of Response	Benchmark Literacy	SAT-9 Reading	Benchmark Mathematics	SAT-9 Mathematics
	Reading Efficacy		Math Efficacy	
Not at all/Never	177	624	143	647
Very Little	195	648	207	650
Sometimes	204	657	186	639
Quite a Bit/Often	208	671	246	667
A Lot/All the Time	211	678	237	659

Table 6

Differences in Student Performance Based on Efficacy Judgements

Question	Test	F	p	Effect Size (Δ)	Means (N)
Q21	Benchmark Math	6.44	.0119	.61	165(19) vs 220(185)
	SAT-9 Math	9.55	.0022	.73	625(20) vs 654(209)
Q23	Benchmark Literacy	12.79	.0004	.79	191(22) vs 209(182)
	SAT-9 Reading	13.71	.0003	.72	640(29) vs 670(201)

* Only those tests with statistically significant values were reported. Questions 22 and 24 had no statistically significant test results.

** All model assumptions where checked, and in places where there were violations only effect sizes were reported.

Figure 1. Percentage of Students Responding in Each Category for Survey Questions

	Not at all Never	Very little	Some-times	Quite a Bit/Often	A lot/all the time
I feel anxious when					
• The teacher is preparing us to take an achievement test, like the SAT-9 test. (Q1)	25	25	29	12	9
• The teacher passes out the test. (Q2)	34	31	20	8	7
• I have difficulty answering test questions. (Q3)	10	25	36	15	14
• The teacher calls "time" before I am finished with the test. (Q4)	34	14	12	17	23
I am anxious about					
• How well I will do on the test. (Q5)	9	15	20	23	34
• How well I will do on the math section of the test. (Q6)	19	19	19	18	26
• How well I will do on the reading section of the test. (Q7)	17	26	22	19	16
• What my parents will think about my scores. (Q8)	16	16	17	16	36
• What my teacher will think about my scores. (Q9)	17	20	24	19	21
• How my scores will compare with other students. (Q10)	28	19	22	14	17
• How my scores will affect my progress in school. (Q11)	13	20	15	16	36
• I feel pressure from my parents to score high on the test. (Q12)	34	23	20	12	11

	Not at all Never	Very little	Some-times	Quite a Bit/Often	A lot/all the time
• I feel pressure from my teachers to score high on the test. (Q13)	27	31	24	9	9
• Teachers talk poorly about students whose scores are low. (Q14)	50	21	17	5	7
• I think the SAT-9 test does a good job of of measuring how much I know. (Q15)	14	9	19	26	32
• I believe that taking the SAT-9 tests motivates me to learn. (Q16)	14	17	20	22	27
• My teacher seems stressed about giving the SAT-9 test. (Q17)	41	28	18	4	9
• Our class receives rewards for scoring high on the SAT-9 test. (Q18)	54	14	22	3	8
• Our class receives rewards if we improve our scores on the SAT-9 test. (Q19)	43	23	18	8	9
• I like the testing week because we have less homework and less instruction in the class. (Q20)	16	7	16	9	51
• I do well in math. (Q21)	7	5	29	26	33
• I enjoy doing math. (Q22)	15	16	20	12	38
• I do well in reading. (Q23) 4	8	24	35	31	
• I enjoy reading. (Q24)	11	11	22	16	40

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM033629

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Impact of accountability and school testing on students: Is there evidence of anxiety?

Author(s): Sean W Mulvenon, Joanie V. Connors & Denise Linares

Corporate Source: Office for Research and Measurement in Education U of A

Publication Date: 11/14/01

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1 ↑ ☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A ↑ ☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B ↑ ☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Joanie V. Connors, Ph.D	Printed Name/Position/Title: Joanie V. Connors, Clinical Asst Professor
Organization/Address: ORME, WAAX, U of A, Fayetteville, AR 72701	Telephone: 501 575-3582 FAX: 501 575-2492
	E-Mail Address: jconnors@uark.edu Date: 11/14/01

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)