Working Paper Series

Appalachian Collaborative Center for Learning, Assessment and Instruction in Mathematics

Instructional Software and "No Child Left Behind" in Poor, Rural Schools: Rationality and the Subversion of Math Achievement

Robert Bickel
Marshall University

Connie Cadle
West Virginia University, Parkersburg

January 2004

ACCLAIM's mission is the cultivation of *indigenous leadership capacity* for the improvement of school mathematics in rural places. The Center addresses the mission through efforts to (1) understand the rural context as it pertains to learning and teaching mathematics; (2) articulate in scholarly works, including empirical research, the meaning and utility of that learning and teaching among, for, and by rural people; and (3) improve the professional development of mathematics teachers and leaders in and for rural communities.

Address: Ohio University
119C McCracken Hall
Athens, OH 45701-2979

Office: 740-593-9869
Fax: 740-593-0477

E-mail: howleyc@ohio.edu
Web: http://acclaim.coe.ohiou.edu

Funded by the National Science Foundation as a Center for Learning and Teaching, ACCLAIM is a partnership of the University of Tennessee (Knoxville), University of Kentucky (Lexington), Kentucky Science and Technology Corporation (Lexington), Marshall University (Huntington, WV), West Virginia University (Morgantown), the University of Louisville, and Ohio University (Athens).

This material is based upon the work supported by the National Science Foundation Under Grant No. 0119679. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Instructional Software and "No Child Left Behind" in Poor, Rural Schools: Rationality and the Subversion of Math Achievement

Robert Bickel

Marshall University

Connie Cadle

West Virginia University, Parkersburg

September 11, 2003

Abstract

Standardized achievement tests as measures of school performance are an inescapable fact of life in U.S. public education. Critics of such tests hold that U.S. education policy has made achievement test scores more important than achievement itself. "No Child Left Behind," the centerpiece of the Bush Administration's domestic agenda, places a premium on standardized testing as a means of reforming public education and raising all students to an acceptable level of achievement. As presented by the U.S. Secretary of Education, the Act promises to be especially helpful in economically disadvantaged, socially devalued, rural areas, such as Appalachian West Virginia. While "No Child Left Behind" nominally gives pride of place to literacy, the Administration has voiced concern that deficiencies in math achievement, especially among economically deprived students in rural schools, threaten the economic pre-eminence of the U.S. This concern was recently manifest in a $22 million grant to the University of Kentucky. Following evaluation of a short-term, non-intensive, in-school intervention called Sophisticated Software, however, we tentatively conclude that improved test-taking skills are easily mistaken for math achievement growth. "No Child Left Behind," thus, may have vastly improved market conditions for the instructional medium that Sophisticated Software typifies. As a result, one consequence of the Act may be to undercut math achievement in poor, rural schools, while providing the appearance of math achievement gains.

Introduction

Standardized tests as measures of school performance are a ubiquitous fact of public life in the U.S. (Bolon, 2000; Dorn, 2003). This is manifest in and reinforced by "No Child Left Behind," the centerpiece of the Bush Administration's domestic policy agenda (U.S. Department of Education, 2002a). The value of such measures in gauging school-engendered achievement is taken for granted by most observers. Nevertheless, pervasive use of standardized tests remains a subject of controversy and critique for a vocal minority (see, for example, Cardenas, 1998; Klein, Hamilton, McCaffrey, & Stecher, 2000; Goodson and Foote, 2001).

Among those most suspicious of high-stakes standardized testing are critical researchers who hold that achievement test scores have taken precedence over achievement (Haney, 2000). In this view, rational responses to indiscriminate use of test-based standards include some teaching explicitly for the test (Grant, 2000; Cimbricz, 2002). Arguably, the rationality of this response becomes more compelling as testing becomes more intrusive, sanctions are tied to test scores, and testing interferes with learning (Camilli and Bulkley, 2001; Mui, 2003). Teaching children to take tests can, hypothetically, minimize the instructional time necessary to fulfill accountability expectations.

In the following, we report results of an evaluation of an instructional software package marketed to enable schools to meet the demands of high-stakes standardized testing programs, most conspicuously those required by "No Child Left Behind." We have given the instructional software a pseudonym, Sophisticated Software. We use a pseudonym because it is the medium that is important, not the brand. Sophisticated Software typifies a medium made more marketable by the increasing intrusiveness of high-stakes standardized tests in public schools (see, for example, Chancery Student Management Solutions, 2001; Curriculum Advantage, 2003; Test Maker, 2003; Pearson Digital Learning, 2003; The Critical Thinking Company, 2003; Merit Software, 2003).

Using results from one poor, rural Appalachian secondary school, our analyses suggest that Sophisticated Software is equally effective in promoting higher test scores in disciplines, in this case math, whether or not students have first undergone an instructional module in the subject. A reasonable, even if tentative, conclusion is that Sophisticated Software does a first-rate job of teaching students to take standardized tests, whatever their content. By dramatically increasing demand for higher standardized test scores, "No Child Left Behind" provides new market-based incentives for the instructional medium which Sophisticated Software typifies. In the process, "No Child Left Behind" may undercut real math achievement in poor, rural schools, giving pride of place to the quantitative appearance of achievement.

"No Child Left Behind"

As with his time as Governor of Texas, President Bush has promulgated policies that tie judgments about school performance to high-stakes standardized tests (Haney, 2002). This is conspicuously evident in the most recent re-authorization of the Elementary and Secondary Education Act of 1965, commonly referred to as "No Child Left Behind."

While the omnipresence of standardized testing is an old story, the federal-level sanctions built into "No Child Left Behind" are novel, indeed. Failure to move students toward mandated performance levels forces schools to invoke a variety of costly correctives. These include providing vouchers to facilitate transfer from poorly performing schools to public alternatives, complemented with supplemental services, including private tutoring (Dillon, 2003).

Predictably, an oft-noted consequence of "No Child Left Behind" is expansion of the role of the federal government in public education (Hardy, 2003). For the first time, federal involvement in local schools and school systems is to be regulated by scores on federally mandated testing programs. The controversial nature of the Act, however, is reflected in the Bush Administration's counter assertion that "No Child Left Behind" actually increases flexibility and control at the local level: what some take to be expansion of federal authority is better construed as redefinition (U.S. Department of Education, 2003).

According to this interpretation, "No Child Left Behind" will enable local educators to identify research-based policies tailored to local needs. As such, the intended effects of the Act broadly parallel those of the rural systemic initiatives funded by the National Science Foundation to promote social and economic development in poor, rural areas through improved math and science education tailored to local opportunities and constraints (National Science Foundation, 1997).

No Child Left Behind in Poor, Rural Areas

A first reading of the Act, moreover, seems to promise higher achievement for victims of persistent poverty and culturally pervasive ascriptions of insularity and reckless indifference, such as those prevailing for rural Appalachia. The primary purpose of the Act, as explained by the U. S. Secretary of Education, is to employ federal funds to close the achievement gap between disadvantaged students and their peers, raising all students to a proficient level (U.S. Department of Education, 2003).

Furthermore, the social psychological rationale for "No Child Left Behind" seems suited to the needs of victims of enduring rural poverty and pernicious Appalachian stereotypes. The

Act is premised on the assumption that effective schools need not be constrained by overarching contextual factors or students' socially ascribed traits. This encouraging judgment is couched in terms of expectations: raise expectations for disadvantaged, socially devalued students, and they will rise to the occasion. Otherwise, students become victims of what the Secretary has termed "the soft bigotry of low expectations" (U.S. Department of Education, 2003).

The "National Math and Science and Education Gap"

Reading achievement has been given most attention in accounts of "No Child Left Behind." The Act's "Reading First" grants provide money targeted directly toward increasing literacy and improving reading achievement (International Reading Association, 2003).

The Bush Administration, however, has voiced strong concern as to the need to improve math achievement to maintain world economic leadership. Although some observers have been dismayed by limited funding for math education (American Society of Mechanical Engineers, 2003; American Institute of Physics, 2003), the University of Kentucky recently received a $22 million grant from the National Science Foundation's Math and Science Partnership Program. Characterized as a "key facet of … 'No Child Left Behind,'" the grant is intended to improve math and science education among poor, rural schools in central Appalachia, and to help schools in the U.S. catch up with schools in other countries (Bozeman, 2002; West Virginia Department of Education, 2003b).

A Dangerous Oversimplification?

"No Child Left Behind" seems disarmingly modern and straightforward, intended to help those with the most pressing educational needs. Scientifically validated methods of promoting achievement in math and a broad range of other disciplines, incorporated with an eye to local requirements and coupled with high expectations for all, will enable poor, rural students to shake off the constraints of context and class.

For many professional educators, however, “No Child Left Behind” dangerously oversimplifies the social circumstances of education (Coles, 2001; Bianchini, 2002; Denlinger, 2002; Huston, 2003). In this view, the effects of context, class, and other non-meritocratic intrusions cannot be avoided with the ease the Act suggests. In addition, the commonly made claim that there are research-based “best practices” that predictably make schools more effective seems dubious to some observers (see, for example, Shafer, 1997; Sarason, 1999; Bickel, Tomasek, and Eagle, 2000; Ross, J., Hannay, L. & Hogaboam-Gray, 2001; Weiner, 2003). Too often, critics have argued, such ostensibly best practices have turned out to be ineffective fads (Pogrow, 2001; Stone, 2002).

Furthermore, while “No Child Left Behind” may seem especially attuned to helping economically disadvantaged, rural schools in culturally devalued places such as Appalachia, resistance to central provisions of the Act has been especially strong in poor, rural areas. There is growing concern among rural educators and officials from rural states that rural contexts have been misunderstood by the authors of “No Child Left Behind” (Tompkins, 2003; Rural School and Community Trust, 2003).

Schooling for “No Child Left Behind”

From its inception, nevertheless, “No Child Left Behind” has enjoyed national-level bipartisan support (Miller, 2003). Whether or not it has *merit*, the Act promises to become a durably important educational *issue* for schools and districts throughout the U.S.

Consequently, among educators who suspect that concern for effective and equitable schooling may be obscured by undue emphasis on test scores, rational responses may include teaching for the test. This makes real achievement gains secondary to increased test scores, whatever they mean (Kaufhold, 1998; Cavalcante, 2001; Haney, 2000). In this way, "No Child Left Behind" creates private profit-making opportunities in public education for entrepreneurs with products that promise test score improvements (Clowes, 2002; Education Industry Association, 2002; Neas, 2003).

Sophisticated Software Products

Sophisticated Software products are marketed to meet the demand produced by the kind of test score-driven sanctions built into "No Child Left Behind." As a personal computer-based intervention, Sophisticated Software offers the following list of "Key Benefits": "Immediate Feedback," Targeted Instruction," "Helps Prepare for Standardized Tests," "Data-Driven Decision Making," "Research Based," "Free Technical Support," and "Try Before You Buy." With regard to preparing students for standardized tests, Sophisticated Software's manufacturer notes that "educators have reported that a student's score on our programs can be an accurate predictor of future performance on state and national exams. We cover many of the same key benchmarks against which students are measured on standardized tests."

Sophisticated Software as a Good-Faith Intervention

For the most part, however, there is little about the way that Sophisticated Software is promoted that suggests that it should be construed as a method for teaching the content of any specific test or test-taking skills generally. The stated emphasis is on real achievement growth.

Recent developments in math, for example, include “two new programs to develop students’ basic algebra thinking skills.” The “Pre-Algebra Shape-up” and “Basic Algebra Shape-up” programs are intended to “help students master … converting metric units, solving simple equations, reading line and bar graphs, performing operations with integers, creating formulas, and using ratios and proportions.”

Aimed at students at grade levels six through nine, the new math programs include step-by-step tutorials with hundreds of questions and problems, immediate responses to student efforts, with answers and worked-out solutions, and summary pretests and posttests. The advertised purpose of Sophisticated Software’s new algebra programs seems plain: to foster real achievement gains, rather than to speciously inflate standardized achievement test scores.

Sophisticated Software in Use: The School and District

Use of Sophisticated Software seems consistent with the claim that the primary purpose of “No Child Left Behind” is to close the achievement gap between economically disadvantaged, socially devalued students and their peers, raising all to a proficient level. Moreover, the school and district in which Sophisticated Software was evaluated typify contexts where “No Child Left Behind” promises the most while instilling the most fear: those serving economically disadvantaged students in socially devalued areas, such as rural Appalachian West Virginia. .

The school is a combined middle school and high school, the only secondary school in the district. There are 373 middle school students in grades five through eight, and 424 high school students in grades nine through twelve. Approximately 98.6 percent of all students are white, and 65.7 percent are sufficiently poor to be eligible for the free- or reduced-price-lunch program. Nevertheless, across the middle school grades, mean basic skills scores (math, reading, and language combined) on the nationally normed Stanford 9 achievement test ranged from the 58^{th} percentile to the 68^{th} percentile. The same range for the high school grades went from the 49^{th} percentile to the 57^{th} percentile (West Virginia Department of Education, 2003a).

In 2001, the district had a population of 7,392 spread across 281 square miles, resulting in a population density of 26.3 people per square mile. The median family income was $26,701, 63.6 percent of the national average, and 74.7 percent of the no-frills West Virginia Self-Sufficiency Standard for "a young family of four" (Tuckwiller, 2002). Of families with children, 28.2 percent had incomes below the federal poverty level in a state where 17.9 percent of all families were below that level; the same figure for the entire U.S. was 12.4 percent. The district is 100.0 percent rural, in a state that is 63.9 percent rural; the same figure for the entire U.S. is 24.8 percent. The county seat is the largest town, with a population of 671 (U.S. Census Bureau, 2003).

Sophisticated Software in Use: Supervised Sessions

In an effort to promote active teacher involvement, all sixth and eighth grade teachers were given an opportunity to participate or not participate in the Sophisticated Software intervention. Two of the teachers who agreed to participate were certified in language arts and reading; they acted as lead teachers, identifying other teachers who would volunteer to participate. One sixth-grade teacher and one eighth-grade teacher elected not to participate but offered to have their students serve as the control group. Participating teachers were trained in use of Sophisticated Software during two day-long training sessions arranged by Sophisticated Software representatives.

The school has a new computer lab with thirty IBM-compatible personal computers. The lab was designed for group instruction, as well as individual computer-assisted instruction. Sophisticated Software was installed on the computers by the school district's Internet technician.

The one hundred sixteen students participating in the intervention were assigned to 12 teacher-supervised groups ranging in size from seven to 14 students. Half of all student participants were in the sixth grade and half were in the eighth grade. The control group, too, had half sixth-grade and half eighth-grade students.

Groups of Sophisticated Software participants were expected to meet for two forty-five minute sessions each week during the Spring 2003 semester. Missed sessions due to illness, snow days, and similar logistical matters were made up whenever possible. At the end of the semester, actual participation ranged from four to 12 sessions, with a mean of 8.10 and a median of 8.00.

The Sophisticated Software resources used by this one poor, rural school provided short, non-fiction selections designed to improve specific reading skills, including identifying the main idea, following sequences of events, vocabulary, factual recall, making inferences, and drawing conclusions. Students were also given opportunities to apply this instruction in writing assignments.

Sophisticated Software enabled teachers to closely monitor student performance and modify instruction accordingly. Participating teachers, thus, served as facilitators and evaluators, answering student questions, assuring that unfamiliar software conventions did not interfere with the learning of substance, and individualizing instruction.

No instructional modules in math were used. The forty-three students in the control group continued with their usual schedules.

Evaluating Sophisticated Software with the Stanford 9

In the following pages, we evaluate Sophisticated Software as a means of improving student achievement, as measured by West Virginia's state-mandated measure of school performance, the widely used Stanford 9 (Case & Slawski, 2003). Conceptually and methodologically simpler applications of the Stanford 9 to study achievement growth are commonplace in applied research and program evaluation (see, for example, Turnbull, Welsh, Held, Davis, & Ratnofsky, 1999; San Juan Unified School District, 2000; Balfanz, Spiridakis & Neild, 2002; Salinas High School, 2002; Kellor, 2003; Gandara and Rumberger, 2003). The Stanford 9 is often the test chosen to gauge school performance in meeting the requirements of "No Child Left Behind" (see, for example, U.S. Department of Education, 2002b; District of Columbia Board of Education, 2003; California State Board of Education, 2003).

We use as outcome measures four sections of the Stanford 9 pertinent to the present application: reading vocabulary, reading comprehension, language mechanics, and language expression.

In addition, even though Sophisticated Software modules in math were not used in our one-school intervention, we also use as outcome measures the Stanford 9 tests in math problem solving and math achievement. For all grades at the middle school level, the math problem-solving test is intended to gauge students' knowledge of number and number relationships, number systems and number theory, algebra, probability, patterns and functions, statistics, geometry, measurement, estimation, and problem solving strategies. The math procedures test is intended to measure students' knowledge of computation using symbolic notation, rounding, and computation in context (West Virginia Department of Education, 2003a).

Why Math?

The choice of the Stanford 9 reading and language tests in evaluating Sophisticated Software is obvious: these four tests gauge student performance in subjects very closely related to what is actually taught with the Sophisticated Software Intervention. The math problem-solving and math procedures sections, however, seem dubious choices, since Sophisticated Software was not used to teach math.

Our objective in nonetheless including math is to provide a tentative answer to the following question: "What does Sophisticated Software teach?" More pointedly, does Sophisticated Software teach subject matter, or does it teach test-taking skills? This question is arguably important in the case of an instructional medium that has arguably become more marketable with the advent of "No Child Left Behind" and the rise of "Annual Yearly Progress" based on (among other things) children's competence in negotiating state-mandated accountability tests.

Decision-Making Rules

If results of our evaluation of Sophisticated Software show that it improves student test scores on none of the sections of the Stanford 9, we will tentatively conclude that the intervention promotes neither achievement growth nor development of test-taking skills.

If results of our evaluation show that Sophisticated Software improves student scores on one or more of the tests in subjects that Sophisticated Software was actually used to teach, but not in math, we will tentatively conclude that the intervention promotes real achievement growth in the subjects it is used to teach.

If the results of our evaluation show that Sophisticated Software improves test scores in subjects it was used to teach, and contributes as much or more to improving scores on math tests, we will tentatively conclude that the intervention primarily teaches test-taking skills.

These three are not the only possible outcomes. For unexplained reasons, Sophisticated Software could be associated with increases in math scores only. In addition, Sophisticated Software may teach both test-taking skills and substantive knowledge. Moreover, the intervention may improve test-taking skills in some subjects but not in others. Because of possibilities such as these, we have phrased our decision-making rules cautiously and emphasized that any conclusions are tentative.

<u>A Quasi-Experimental Assessment of Achievement Growth</u>

Within our poor, rural school, practical considerations prevented random assignment of the 159 sixth and eighth graders to treatment and control groups. As a result, our quasi-experimental design uses judiciously selected control variables to deal with confounding of the Sophisticated Software treatment with extraneous factors.

Furthermore, since we are interested in achievement growth over time, we have used the repeated measures procedure available with SPSS 11.5 Mixed Models. This accommodates unbalanced designs in which the number of cases per group varies, and the total number of cases is not constant across all observations. Growth curves are estimated using the information available, including instances in which test-takers skip an administration (cf. Bryk and Raudenbush, 2002).

In addition, the repeated measures procedure permits modeling of the error term to correct coefficient standard error estimates for the effects of heteroscedasticity and autocorrelation. In the analyses reported below, autocorrelation was not present, but the error terms were modeled to accommodate departures from homoscedasticity, thereby avoiding inflated standard errors and loss of statistical power.

Multi-Level Analysis

Our quasi-experimental growth model uses independent variables measured at three levels: within subjects for repeated measures; between subjects; and between groups, for our twelve teacher-supervised groups. All variables are described in Table 1.

TABLE 1

VARIABLES

INDEPENDENT

Gender	Coded 1 if Male and 0 if Female
GPA	Grade Point Average at End of 2002-2003 School Year
Repeated Grades	Number of Grades Student Repeated
Ethnicity	Coded 1 if White and 0 Otherwise
Free Lunch	Coded 1 if Eligible for Free/Reduced-Cost Lunch and 0 Otherwise
Enrolled Late	Code 1 if Enrolled More than Six Weeks Late and 0 Otherwise
Special	Coded 1 if Designated Special Education and 0 Otherwise
Time	Coded 0 for Pretest and 1 for Posttest Over One Calendar Year
Personal Computer	Coded 1 if PC in the Home and 0 Otherwise
Grade	Either Sixth Grade or Eighth Grade
Sophisticated Software Group	Coded 1 for Participants and 0 Otherwise
Sophisticated Software Sessions	Number of Sessions in which Student Participated

DEPENDENT

(Selected from Stanford 9 Battery)

PRE-TESTS (End of Fifth/Seventh Grade)	POST-TESTS (End of Sixth/Eighth)
Reading Vocabulary	Reading Vocabulary
Reading Comprehension	Reading Comprehension
Language Mechanics	Language Mechanics
Language Expression	Language Expression
Math Problem Solving	Math Problem Solving
Math Procedures	Math Procedures

Time is a within-subjects independent variable. Gender, GPA, Repeated Grades, Ethnicity, Free Lunch, Enrolled Late, Special, Personal Computer, and Sophisticated Software Sessions are between-subjects independent variables. Grade and Sophisticated Software Group are between-groups independent variables. Though the teacher- supervised groups varied in size, group size was not used as an independent variable because it applies only to those who participated in the Sophisticated Software intervention, not to those who collectively provided a basis for comparison.

Variables of special interest are Sophisticated Software Group and Sophisticated Software Sessions. As explained in the Results section, these are the variables that enable us to tentatively judge the consequences of using Sophisticated Software in preparation for taking high-stakes standardized achievement tests, such as the Stanford 9.

<u>Why Multiple Levels?</u>

Use of multiple levels in our growth model reflects the fact that intercepts and slopes corresponding to relationships of interest may vary within students in a repeated-measures analysis, and between students within the teacher-supervised groups and the control group. This set of circumstances may be effectively addressed by using random intercepts and random slopes, with their variability treated as function of higher-level independent variables (Longford, 1994).

For the analyses reported below, however, all intra-class correlations between achievement growth and the between-student level were statistically non-significant. Moreover, while the intra-class correlations between achievement growth and the within-student level were statistically significant, none of the slopes corresponding to relationships between the within-student independent variable Time and our Stanford 9 outcome measures had a statistically significant variance, though the intercept variance was statistically significant for each Stanford 9 outcome measure.

As a result, the only parameters permitted to randomly vary are the intercepts corresponding to relationships between achievement and the within-student level. The random intercepts, thus, are specified as functions of the between-student independent variables, while the remaining intercepts and all slopes are fixed. The equation for our one random intercept is as follows:

$$\beta_{0J0} = \gamma_{000} + \gamma_{010}\mathrm{GENDER} + \gamma_{020}\mathrm{GPA} + \gamma_{030}\mathrm{REPEAT} + \gamma_{040}\mathrm{ETHNIC}$$
$$+ \gamma_{050}\mathrm{LUNCH} + \gamma_{060}\mathrm{EARLY} + \gamma_{070}\mathrm{SPECIAL} + \gamma_{080}\mathrm{PC} + u_{0J0}$$

Results

Descriptive statistics computed before centering with respect to grand means for between-groups independent variables and with respect to teacher-group means for between-student independent variables appear in Table 2.

TABLE 2

DESCRIPTIVE STATISTICS

Variable	Mean	Standard Deviation	Minimum	Maximum
Gender	0.53	0.50	0	1
GPA	2.74	0.76	0.00	4.00
Repeated Grades	0.37	0.68	0	4
Ethnicity	0.98	0.12	0	1
Free Lunch	0.58	0.49	0	1
Enrolled Late	0.69	0.25	0	1
Special	0.08	0.27	0	1
Time	0.51	0.50	0	1
Personal Computer	0.74	0.44	0	1
Grade	7.01	1.00	6	8
Soph. Software Group	0.82	0.38	0	1
Soph. Software Sessions	5.96	3.62	0	12
Reading Vocabulary	20.72	5.03	5.00	30.00
Reading Comprehension	35.04	9.24	3.00	52.00
Language Mechanics	15.59	4.45	3.00	24.00
Language Expression	15.11	4.79	3.00	24.00
Math Problem Solving	32.41	7.94	10.00	49.00
Math Procedures	18.62	6.37	3.00	30.00

Tables 3 through 6 report regression analysis results for dependent variables corresponding to subjects that Sophisticated Software was intended to teach in this poor, rural school in Appalachian West Virginia: reading vocabulary, reading comprehension, language mechanics, and language expression. Tables 7 and 8 report regression results for dependent variables corresponding to subjects Sophisticated Software was not intended to teach in this application. For all tables, unstandardized regression coefficients are estimated using restricted maximum likelihood (REML) estimators (Heck and Thomas, 2002). McFadden's R^2_L "goodness of fit" measure is computed as shown in each table, and can be interpreted as analogous to the conventional R^2 statistic used with ordinary least squares multiple regression analysis (Menard, 2002).

Sophisticated Software Group and Sophisticated Software Sessions

Our primary interest is in coefficients corresponding to the Sophisticated Software Group and Sophisticated Software Sessions independent variables. Sophisticated Software Group is coded 1 for students who used Sophisticated Software, and 0 for those who did not. Sophisticated Software Sessions refers to the number of 45-minute instructional sessions in which students actually participated.

Typically, we would expect variables created in this way to be closely correlated. In this instance, however, the Sophisticated Software Group and Sophisticated Software Sessions variables are orthogonal. This is because Sophisticated Software Sessions is best construed as an interaction effect computed by multiplying the Group dummy variable by the number of Sessions in which a student actually participated. Since the Group variable is coded 0 for Sophisticated Software non-participants, these products will be equal to zero. Since the Group variable is coded 1 for Sophisticated Software participants, these products will be equal to the value of Sophisticated Software Sessions itself. Finally, given that Sophisticated Software Group is centered with respect to its grand mean and Sophisticated Software Sessions is centered with respect to group means, the interaction effect created using these variables will be orthogonal to them (Iversen, 1991).

The foregoing discussion of the meaning of the Sophisticated Software Sessions variable also explains why we have not used a Group-by-Sessions multiplicative interaction term: with Sessions already in the regression equation, adding the interaction term is redundant. Even with centering, the bivariate correlation between Sessions and Group-by-Sessions is .75, suggesting serious multicollinearity problems.

The coefficients corresponding to Sophisticated Software Group and Sophisticated Software Sessions, therefore, are interpreted as follows: the numerical magnitude of a statistically significant and positive Group coefficient tells us how much, on average, a student's achievement growth is increased simply by virtue of the fact that he or she participates in the Sophisticated Software intervention. The numerical magnitude of a statistically significant and positive coefficient corresponding to Sessions tells us how much additional achievement growth, on average, comes with actual participation in an instructional session.

Regression Results for Subjects Taught

Tables 3 and 4 show us that the Sophisticated Software Group variable has statistically significant and positive regression coefficients when the Stanford 9 reading vocabulary and reading comprehension tests are used as dependent variables. Participation in the Sophisticated Software Group increases achievement growth in reading vocabulary, on average, by 2.54 points, while participation in the Sophisticated Software Group increases achievement growth in reading comprehension, on average, by 3.47 points. (The reported sample size reported for each analysis is less than the original 159 cases due to missing data on one or more variables.)

TABLE 3

STANFORD-9 READING VOCABULARY

Parameter	Parameter Estimate	t Value	Sig. Level*
Intercept	***20.59***	***64.84***	***.000***
Gender	***2.16***	***3.06***	***.000***
GPA	***3.10***	***6.16***	***.000***
Repeated Grades	- 0.44	-0.85	.297
Ethnicity	***- 5.74***	***-2.10***	***.018***
Free Lunch	- 1.12	-1.49	.065
Enrolled Late	1.20	0.88	.184
Special Education	***-3.56***	***-2.94***	***.002***
Time	***1.41***	***3.97***	***.000***
Personal Computer	0.01	0.01	.484
Grade	0.05	0.13	.401
Sophisticated Group	***2.54***	***2.60***	***.003***
Sophisticated Sessions	0.06	0.61	.248

* Significance Levels for One-Tailed Tests; bold italics indicates statistically significant variables at $p < .05$

N = 132

$R^2_L = 1693.4 – 1353.4/1693.4 = 20.1\%$

TABLE 4

STANFORD-9 READING COMPREHENSION

Parameter	Parameter Estimate	t Value	Sig. Level*
Intercept	***34.48***	***62.47***	***.000***
Gender	***2.12***	***1.76***	***.039***
GPA	***7.39***	***8.59***	***.000***
Repeated Grades	- 0.93	- 1.04	.152
Ethnicity	***- 9.70***	***- 2.06***	***.021***
Free Lunch	- 0.84	- 0.65	.251
Enrolled Late	2.54	1.05	.147
Special Education	- 2.74	-1.32	.093
Time	1.11	1.63	.052
Personal Computer	- 0.80	- 0.57	.291
Grade	0.84	1.30	.097
Sophisticated Group	***3.47***	***2.07***	***.019***
Sophisticated Sessions	0.12	0.70	.307

* Significance Levels for One-Tailed Tests; bold italics indicates statistically significant variables at $p < .05$

N = 132

R^2_L = 2059.0 – 1659.7/2059.0 = 19.4%

Table 5 shows us that, while the coefficient corresponding to Sophisticated Group is not statistically significant with the Stanford 9 Language Mechanics score as the dependent variable, the coefficient corresponding to Sophisticated Software Sessions is statistically significant and positive. For each instructional session in which a student actually participated, the achievement gain equaled, on average, 0.21 points.

TABLE 5

STANFORD-9 LANGUAGE MECHANICS

Parameter	Parameter Estimate	t Value	Sig. Level*
Intercept	***15.46***	***54.36***	***.000***
Gender	0.42	0.65	.258
GPA	***3.11***	***6.94***	***.000***
Repeated Grades	***- 0.80***	***- 1.71***	***.047***
Ethnicity	- 0.44	- 0.18	.433
Free Lunch	- 1.07	- 1.58	.057
Enrolled Late	0.34	0.28	.390
Special Education	***-2.76***	***- 2.55***	***.006***
Time	***0.67***	***2.18***	***.015***
Personal Computer	- 0.78	- 1.05	.147
Grade	- 0.18	- 0.54	.295
Sophisticated Group	0.83	0.95	.171
Sophisticated Sessions	***0.21***	***2.34***	***.010***

* Significance Levels for One-Tailed Tests; bold italics indicates statistically significant variables at $p < .05$

N = 132

$R^2_L = 1644.1 – 1308.4/1644.1 = 20.4\%$

In Table 6 we see that, with the Stanford 9 Language Expression test score as the dependent variable, both Sophisticated Software Group and Sophisticated Software Sessions have coefficients which are statistically significant and positive. Membership in the Sophisticated Software Group yields an achievement growth increase equal, on average, to 1.75 points. Furthermore, each Sophisticated Software Session in which students actually participate yields an additional achievement growth increase equal, on average, to 0.19 points.

TABLE 6

STANFORD-9 LANGUAGE EXPRESSION

Parameter	Parameter Estimate	t Value	Sig. Level*
Intercept	***14.77***	***47.77***	***.000***
Gender	0.94	1.37	.085
GPA	***3.61***	***7.40***	***.000***
Repeated Grades	- 0.49	- 0.97	.167
Ethnicity	- 4.01	- 1.50	.067
Free Lunch	0.71	- 0.96	.169
Enrolled Late	1.66	1.24	.108
Special Education	***- 2.36***	***- 2.00***	***.023***
Time	***0.88***	***2.58***	***.005***
Personal Computer	- 1.12	- 1.41	.079
Grade	0.05	0.15	.440
Sophisticated Group	***1.75***	***1.85***	***.032***
Sophisticated Sessions	***0.19***	***1.86***	***.031***

* Significance Levels for One-Tailed Tests; bold italics indicates statistically significant variables at $p < .05$

$N = 132$

$R^2_L = 1692.0 - 1356.7/1692.0 = 19.8\%$

In view of the brevity and limited intensity of the Sophisticated Software intervention, these achievement gains are quite large, and they raise an obvious question: does Sophisticated Software teach difficult content, or does it teach easier-to-learn standardized test-taking skills?

Regression Results for Math

Tables 7 and 8 report regression results for dependent variables for which Sophisticated Software modules are produced but were ***not*** used in this poor, rural school application: math problem solving and math procedures. Both analyses yield statistically significant and positive regression coefficients corresponding to Sophisticated Software Group and Sophisticated Software Sessions. With the Stanford 9 math problem-solving test as the dependent variable, membership in the Sophisticated Software Group yields an increase in achievement growth equal, on average, to 2.58 points, while participation in each Sophisticated Software Session yields an additional 0.36 points. Similarly, with the math procedures test as the dependent variable, membership in the Sophisticated Software Group yields an increase in achievement growth equal, on average, to 2.71 points, while participation in each Sophisticated Software Session yields an additional 0.38 points.

TABLE 7

STANFORD-9 MATH PROBLEM SOLVING

Parameter	Parameter Estimate	t Value	Sig. Level*
Intercept	***32.10***	***65.79***	***.000***
Gender	***4.36***	***4.05***	***.000***
GPA	***6.54***	***8.57***	***.000***
Repeated Grades	- 1.12	- 1.41	.079
Ethnicity	***- 11.03***	***- 2.63***	***.004***
Free Lunch	- 1.00	- 0.87	.192
Enrolled Late	-0.08	- 0.04	.484
Special Education	***- 4.44***	***- 2.39***	***.008***
Time	***1.03***	***2.27***	***.012***
Personal Computer	- 0.23	- 0.18	.429
Grade	***- 1.39***	***- 2.45***	***.008***
Sophisticated Group	***2.58***	***1.74***	***.041***
Sophisticated Sessions	***0.36***	***2.32***	***.010***

* Significance Levels for One-Tailed Tests; bold italics indicates statistically significant variables at $p < .05$

N = 132

R^2_L = 1968.2 – 1531.0/1968.2 = 22.2%

TABLE 8

STANFORD-9 MATH PROCEDURE

Parameter	Parameter Estimate	t Value	Sig. Level*
Intercept	***18.66***	***50.29***	***.000***
Gender	***2.66***	***3.22***	***.000***
GPA	***4.25***	***7.26***	***.000***
Repeated Grades	***- 1.55***	***- 2.54***	***.006***
Ethnicity	- 1.39	- 0.43	.334
Free Lunch	0.37	0.41	.341
Enrolled Late	1.53	0.95	.171
Special Education	- 1.03	- 0.71	.236
Time	***0.86***	***2.07***	***.015***
Personal Computer	- 1.39	- 1.45	.074
Grade	***- 2.93***	***- 6.72***	***.000***
Sophisticated Group	***2.71***	***2.38***	***.009***
Sophisticated Sessions	***0.38***	***3.21***	***.000***

* Significance Levels for One-Tailed tests; bold italics indicates statistically significant variables at $p < .05$

N = 132

$R^2_L = 1849.6 – 1447.8/1849.6 = 21.7\%$

As with reading vocabulary, reading comprehension, language mechanics, and language expression, given the brevity and limited intensity of the Sophisticated Software intervention, the achievement gains in math problem solving and math procedures seem quite large. In this instance, however, the gains are better characterized as mystifying, because they pertain to subjects not taught in our Appalachian application of Sophisticated Software.

Rural Math Achievement and "No Child Left Behind"

For decades, critics of indiscriminate use of standardized testing have asked "what do standardized achievement tests really measure?" (see, for example, Madaus, 1985; Rotberg, 1995; Klein, Hamilton, McCaffrey, & Stecher, 2000; Bolon, 2001). In addition to achievement, answers have included social class (Carnevale & Rose, 2003), race (Whitworth Communications, 2001); gender (Foley & Redd, 2003), group composition (Barr and Dreeben, 1983), speed (Bolon, 2001), quick but shallow thinking (National Center for Fair & Open Testing, 2003), and the specific form of intelligence needed to be a good psychometrician (Aaron and MacClaury, 1978). There may be merit to each of these answers.

In our evaluation of the application of Sophisticated Software in one poor, rural secondary school in Appalachian West Virginia, however, the answer seems to be *test-taking skills*. These skills, moreover, appear not to be naturally occurring knacks or informally acquired manifestations of cultural capital advantage. Instead, they are taught by short-term, non-intensive use of Sophisticated Software.

When Sophisticated Software modules designed to teach reading and language are used, measured achievement growth in those subjects is increased, but so is measured achievement growth in math. The contribution of Sophisticated Software to measured achievement growth in math problem solving and math procedures is at least as great as the measured contribution to reading vocabulary, reading comprehension, language mechanics, and language expression.

It may be that instruction in reading and language benefits math test-takers because most of the math problem solving and math procedures problems on the Stanford 9 are word problems. Perhaps this provides part of the answer to "What is Taught?" and to "What is

Measured?" Nevertheless, the rest of the answer to both questions seems clearly to be "Not Math!"

Nevertheless, if Sophisticated Software reliably increases measured math achievement, whatever that may mean, "No Child Left Behind" contributes to making it and the medium it represents much more marketable. When damaging sanctions become tied to standardized achievement test scores, rational educators and public policymakers have little choice but to comply. If compliance is facilitated by teaching for test-taking with relatively non-intrusive, comparatively inexpensive, easy-to-use tools such as Sophisticated Software, it is likely that rational participants in the world of education will use them.

When faced with meeting performance standards set by "No Child Left Behind" or else being subject to any of a number of sanctions—paying for students to transfer to other schools, paying for supplemental services from outside providers, replacing staff members, implementing a new curriculum, granting administrative authority to outside experts, extending the school day, or re-opening as a charter school—schools and districts will find ways to meet the standards (Boylan, 2003). With limited resources, that may very well mean placing the emphasis on measures of math achievement rather than math achievement itself. "No Child Left Behind" thereby gives Sophisticated Software and the medium it typifies far more lucrative opportunities than the public schools offered before.

References

Aaron, H. and MacLaury, B. (1978). *Politics and the Professors.* Washington, D.C.: Brookings.

American Institute of Physics (2003). Higher funding sought for math and science partnerships. *The AIP Bulletin of Science and Policy News*. http://www.aip.org/enews/fyi/2003/059.html

American Society of Mechanical Engineers (2002). *Action Alert: Senate to Determine Math and Science Funding*. http://www.asme.org/gric/ActionAlerts/MSPFunding_090502.html

Balfanz, R., Spiridakis, K., and Neild, R. (2002). Will converting high-poverty middle Schools to k-8 schools facilitate achievement gains? A research brief for the school district of Philadelphia. Philadelphia: Philadelphia Education Fund.

Barr, R. and Dreeben, R. (1983). *How Schools Work*. Chicago: University of Chicago.

Bianchini, L. (2002). NCTE resolution on the reading first initiative. *News from NCTE*. http://NCTE.org/news/resolution.shtml

Bickel, R., Tomasek, T., & Eagle, T. (2000). Top-down routinized reform in low-income rural schools. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v8n12.html

Bolon, C. (2000). School-based standard testing. *Education Policy Analysis Archives*. http://epaa.asu/edu/epaa/v8n3

Bolon, C. (2001). Significance of test-based ratings for metropolitan Boston schools. *Education Policy Analysis Archives*. http://epaa.asu.edu/v9n42.html

Boylan, D. (2003). *ESEA and HR1, the No Child Left Behind Act*. Columbus, OH: Buckeye Association of School Administrators.

Bozeman, K. (2002). UK receives $22 million grant to enhance math and science education. Lexington, KY: University of Kentucky, University of Kentucky Public Relations.

Bryk, A. and Raudenbush, S. (2002). *Hierarchical Linear Models*. Thousand Oaks, CA: Sage.

California State Board of Education (2003) NCLB: SBE approves plan for adequate yearly progress. Sacramento, CA: California State Board of Education.

Camilli, G. and Bulkely, K. (2001) Critique of "An evaluation of the Florida A-Plus accountability and school choice program." *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v9n7/

Carnavale, A. and Rose, S. (2003). *Socioeconomic Status, Race/Ethnicity, and Selective College Admissions*. New York: The Century Foundation.

Cavalcante, J. (2001). New, tougher exams reduce stigma of 'teaching to the test.' *The Business Review*. http://www.bizjournals.com/albany/stories/2001/04/16/focus8.html

Chancery Student Management Solutions (2001). Chancery Software adds new level of flexibility to its web-based district-wide student management solutions. http://www.chancery.com/corporate/newsroom/pr/releases.html

Cimbrica, S. (2002). State-mandated testing and teachers' beliefs and practices. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v10n2.html

Clowes, G. (2002). Edventures 2002 teems with entrepreneurs, ideas, excitement. Chicago: Heartland Institute.

Coles, G. (2001). Learning to read "scientifically." *Rethinking Schools Online*. http://www.rethinkingschools.org/archives/15_04/read154.htm

Critical Thinking Company (2003). Discover the power of critical thinking. http://www.criticalthinking.com/index.html

Curriculum Advantage (2003). Classworks: Partnering with schools to make a difference. http://www.classworks.com/pdf/cw_white_paper_nclb.pdf

Denlinger, S. (2002). Teaching as a profession: A look at the problem of teacher deficits. *Clearing House*. *5*: 116-117.

Dillon, S. (2003). Cuts put schools and law to the test. *New York Times*, August 31. Retrieved December 25, 2003 from: http://college4.nytimes.com/guests/articles/2003/08/31/2256354.xml

District of Columbia Department of Education (2003). *Report of the Committee on Teaching and Learning*. Washington, D.C. District of Columbia Department of Education.

Dorn, S. (2003). High-stakes testing and the history of graduation. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v11n1/

Education Industry Association (2002). Behind the new law: Entrepreneurs grab the chance to build careers, help schools. http://www.educationindustry.org/news/archive.php

Foley, R. and Redd, N. (2003). *The Girls Guide to the SAT*. New York: Random House.

Gandara, P. and Rumberger, R. (2003). The inequitable treatment of English learners in California's public schools. Los Angeles: Institute for Democracy, Education, and Access.

Goodson, I. and Foote, M. (2001). Testing times: A case study. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v9n2.html

Grant, S. (2000). Teachers and tests: Exploring teachers' perceptions of changes in the New York State testing program. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v8n14.html

Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v8n41/

Haney, W. (2002). Lake Woebeguaranteed: Misuse of test scores in Massachusetts, Part I. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v10n24/htm

Case, B. and Slawski, E. (2003). *Assessment Report: Accommodations for the Stanford 9*. New York: Harcourt.

Hardy, L. (2003). Overburdened and underwhelmed. ASBJ.com. http://www.sdbj.com/current/coverstory.html

Heck, R. and Thomas, S. (2000). *An Introduction to Multilevel Modeling Techniques*. Mahwah, NJ: Lawrence Erlbaum.

Huston, P. (2003). The bigotry of expectations. *School Administrator*. January. http://www.aasa.org.publications/sa/2003_01/exceper.htm

International Reading Association (2003). IRA survey examines process for Reading First applications. *Reading Today*. February/March.

Iversen, G. (1991). *Contextual Analysis*. Newbury Park, CA: Sage.

Kaufhold, J. (1998) What's wrong with teaching for the test? *The School Administrator* Web Edition, December. www.aasa.org./publications/sa/1998_12/kaufhold.htm

Kellor, E. (2003). Catching Up with the Vaughn express: Four years of performance pay and standards-based teacher evaluation. Madison, WI: Consortium for Policy Research in Education.

Klein, S., Hamilton, L., McCaffrey, D., and Stecher, B. (2000). What do test scores in Texas tell us? *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v8n49

Longford, N. (1993). *Random Coefficient Models*. New York: Oxford.

Madaus, G. (1985) Test scores as administrative mechanisms in educational policy. *Phi Delta Kappan*. *66*: 611-617.

Menard, S. (2002). *Applied Logistic Regression Analysis*. Newbury Park, CA: Sage.

Merit Software (2003). Learn how Merit can help you. http://www.meritsoftware.com/

Miller, G. (2003). Statement by the Honorable George Miller: Press conference on NCLB one-year anniversary shortfall. Washington, D.C.: Committee on Education and the Workforce, U.S. House of Representatives.

Mui, Y (2003). Poor, Others Fall Short on New Maryland Test. *The Washington Post*, August 22. http://www.washingtonpost.com/wp-dyn/articles/A32119-2003Aug22.html

National Center for Fair & Open Testing (2003). What's wrong with standardized tests. http://www.fairtest.org/facts/whatwron.htm

National Science Foundation (1997). Rural systemic initiatives in science, math, and technology education. Washington, D.C.: National Science Foundation, Directorate for Education and Human Resources.

Neas, R. (2003). The voucher veneer: The deeper agenda to privatize public education.Washington, D.C.: People for the American Way Foundation.

Pearson Digital Learning (2003). Meet the challenges of NCLB with mission-critical solutions from Pearson Digital Learning. http://nclb.pearsondigital.com

Pogrow, S. (2001). Success for All never had a research base and never worked. *The Teachers Net Gazette*. http://teachers.net/gazette/MAY01/pogrow.html

Ross, J., Hannay, L., and Hogaboam-Gray, A. (2001) Research in Ontario secondary schools. Toronto: Ontario Ministry of Education and Training and the Social Sciences and Humanities Research Council.

Rotberg, I. (1995). Myths about test score comparisons. *Science, 270*: 607-610.

Rural School and Community Trust (2003). Highly qualified teachers: A primer on NoChild Left Behind. *Rural Policy Matters, 5*: 1.

Salinas High School (2002). School accountability report card 2001-2002. Salinas, CA: Salinas Union High School District.

San Juan Unified School District (2002). Effects of the early reading program on reading achievement. http://www.sanjuan.edu/services/r&e/brief6earlylate.htm

Sarason, S. (1999). Charters: The next flawed reform. *School Administrator, 56*: 32-34.

Shafer, G. (1997). *School Reform and High School Proficiency*. Urbana, IL: National Council of Teachers of English.

Stone, J. (2002). Why NBPTS and other innovations receive so little critical attention. Johnson City, TN: Education Consumers Clearinghouse.

Test Maker (2003). Practice does make perfect. TESTMAKER Series. http://www.ppsisoft.com

Tompkins, R. (2003). Leaving rural children behind. *Education Week* on the Web, March 26 http://www.edweek.org/ew/ewstory.cfm

Tuckwiller, T. (2002). The real cost of living. *Charleston* (WV) *Gazette*, May 26, pp.1A, 6A.

Turnbull, B., Welsh, M., Heid, C., Davis, W., and Ratnofsky, A. (1999). *The Longitudinal Evaluation of School Change and Performance (LESCP) in Title I Schools*. Washington, D.C.: Office of the Undersecretary, U.S. Department of Education.

U.S. Census Bureau (2003). West Virginia QuickFacts from the Census Bureau. http://quickfacts.census.gov/qfd/atates/54000.html

U.S. Department of Education (2002a). Inside No Child Left Behind. My.ED.gov. http://www.ed.gov/legislation/ESEA02/

U.S. Department of Education (2002b). The Achiever. http://www.nclb.glv/Newsletter/20021101.html

U.S. Department of Education (2003). Comments by Secretary Paige to the Commonwealth Club of California. March 12. http://www.ed.gov/02-2003/03122003a.html

Weiner, L. (2003) Research or cheerleading? Scholarship on Community School District 2, New York City. *Education Policy Analysis Archives*. http://epaa.asu.edu/epaa/v11n27/

West Virginia Department of Education (2003a). Data on demand. http://wvde.state.wv.us/

West Virginia Department of Education (2003b). Mathematics education reform initiative for teachers. Charleston, WV: West Virginia Department of Education.

Whitworth Communications (2001). Whitworth professor calls for education reform to address equity as well as accountability standards. Spokane, WA: Whitworth Communications.