

A Study of the Relationship Between Teaching and Context Variables
in Student Outcomes

Shawn M. Quilter

Eastern Michigan University

Results presented at the Fall 2001 Conference of The Renaissance Group
September 27, 2001
Washington, D.C.

Project funded by The Renaissance Group's Research Fellowships Program

A Study of the Relationship Between Teaching and Context Variables in Student Outcomes

Nearly \$12 Billion is spent each year on public education in the United States of America. For various political and social reasons, schools and teachers are increasingly being held accountable for “producing” specific results with this money. Perhaps the most important results are student-centered outcomes. Presumably, the public wants to know that schools are using public monies to effect positive results in all students’ lives.

Student outcomes typically come in the forms of social, emotional, physical, or cognitive development. The most common, and perhaps most publicized measures, are those that reflect cognitive achievements. For the purpose of large-scale accountability, cognitive achievement has traditionally been measured via the results of standardized achievement tests. These tests are typically intended to give a snapshot of student achievement at a particular point in time. Whether commercially produced norm-referenced tests or state-mandated criterion-referenced tests, these assessments, which can include selected and constructed response questions, are commonly used as indicators of achievement. Though there is considerable debate as to the quality and validity of standardized achievement test results for various educational and policy purposes (Wiggins, 1992), standardized achievement test results are a mainstay of evaluation in the American educational system (Barton & Coley, 1994; Ryan, 1995).

There are many different ways that the public can get information about student outcomes in general and standardized achievement tests in particular. Given the current public interest in education, it is not surprising to find that some achievement information is readily available in local newspapers. It is not uncommon today to see achievement test results, especially for state-mandated standards-based tests, listed in newspapers (Olson, 1998). Aggregate test results for local districts and buildings, and even state level results, are typically published after they have been tabulated each year. The publication of raw scores in this manner, however, is problematic at best. Even if we can assume that all of the test data are accurate, which is not always the case, raw test scores are often presented to the public in a decontextualized manner. That is, there is little if any information about the student population, communities served in each district, or the learning context of the school itself. Recently, some newspapers have gotten in the business of trying to provide a context for the test scores they so eagerly report (Olson, 1998). Some papers report demographic information to provide readers with information about the student population of each district in order to put the scores in context. Over the last several years the Detroit Free Press, for example, has conducted regression analyses to predict district level achievement on the Michigan Educational Assessment Program (MEAP) assessments, based on the demographic characteristics of the communities from which the scores came (Olson, 1998). At face value this attempt to provide context seems laudable. However, the regression analyses conducted by many newspapers are so flawed, that the results can be just as misleading if not more misleading than providing raw scores alone, as has been past practice. Besides the potential to use flawed data and inappropriate statistical techniques, newspapers usually use aggregated data for their analyses. The problem with using aggregated data is that it decontextualizes student

achievement because the unit of observation is the school or district, not the student (Haney, 1980). Aggregate level analyses (Bidwell & Kasarda, 1980), whether at the district or building level, are problematic because they tend to gloss over individual student differences within the learning context. Even when newspapers use good data and appropriate techniques for aggregated data, the interpretations made from these results are of limited value since we educate students, not buildings. Thus, it is important that student achievement be analyzed at the student level whenever possible. Class size research is an example of how context and teaching variables were taken into account when considering student outcomes -- specifically cognitive achievement (Finn & Achilles, 1999). The best research in this area is based on student level data and takes teaching and learning context variables into account (Nye, Hedges, Konstantopoulos, 1999).

The main purpose of the current paper is to explore the relationships between teaching/learning context variables and student outcomes, where data are analyzed at the student level. Another purpose of the current paper is to look at the relationship between credentialing from Eastern Michigan University (EMU) and student level outcomes. All analyses are conducted at the student level for this research project.

Following is the research question for this study.

How are teacher and learning context variables related to student achievement test results?

Variables in the Study

Teaching variables included years of teaching experience, professional level (e.g., bachelor, master, etc.), and association with Eastern Michigan University (e.g., degree or credential obtained from EMU versus no degree or credential obtained from EMU). Learning context variables included special needs status, gender, and average attendance (grades 2-4). For this study, the achievement test of interest is the Michigan Educational Assessment Program's (MEAP) essential skills tests in 4th grade: mathematics and reading assessments. Socio-economic status was not included as a variable in this study since fewer than 5% of students qualified for free or reduced lunch.

Methods

Participants

Participants for this study were a convenience sample of all 4th graders at a school district in southeastern Michigan. The school district enrolls approximately 3000 students, of which, approximately 230 were in the 4th grade during the 2000-2001 school year. Approximately 5% of students in the district take free or reduced lunches and the district spends approximately \$7000 per pupil. Most of the children in the sample come from white working class families in a rural/suburban setting. Ethnicity was not included as a variable in this study because fewer than 5% of students belong to an ethnic minority.

Instruments

The only instruments used in this research were the Michigan Educational Assessment Program (MEAP) essential skills tests in mathematics and reading for 4th graders. These tests are state-mandated standards-based tests that Michigan students take in the winter of their 4th grade year (see <http://www.meritaward.state.mi.us/> for details) and during subsequent years in various subject areas. These criterion-referenced tests are used for measuring the extent to which Michigan schools are successfully implementing the state's core curriculum. Both the mathematics and reading tests are multiple-choice tests, though the mathematics test will have a few open-ended questions beginning this year (2001-2002). The reading test is based on students' responses to questions about an informational reading section and a story reading section. Both tests are reported to have adequate reliability of results and content validity for the purpose of estimating the extent to which school districts are successfully implementing the state's core curriculum in each subject area (see <http://www.meritaward.state.mi.us/merit/meap/design.htm>).

The MEAP mathematics and reading assessment results are reported as both a scale score and a rating. Students are rated as either "Satisfactory," "Moderate," or "Low" based on their scale score. "Satisfactory" is an indication that the student has demonstrated mastery of the essential skills for that grade level, based on the core curriculum. "Moderate" is an indication of partial mastery, while "Low" is an indication of non-mastery.

Procedure

Student data were collected from existing student records, known as CA-60s. Information about teacher credentials and the free/reduced lunch status of each student was obtained from each principal. All three schools in the district were visited over the course of several weeks to collect the data.

Once the data were collected, they were coded in SPSS files and analyzed using SPSS version 10.0. Since the research question focused on the relationship between teacher, student, and achievement test results, multiple regression analyses were used to analyze the data.

Limitations

There are at least four limitations that must be considered when interpreting the results of this study. First, it is important to make note of the demographic characteristics of the sample of students and teachers used in this study. The results are probably generalizable only to those students, teachers, and districts that are similar to the sample herein. Second, the optimal procedure for estimating various effects and relationships in this kind of research is nested designs, typically requiring some kind of linear modeling (Bryk & Raudenbush, 1992). Given that the sample included approximately 200 participants, the less complicated regression models that were employed in this study, seemed reasonable for the purpose. Third, while standardized achievement test results are used as the

criterion variable in these analyses, the author recognizes that operationally defining cognitive development in this manner is limiting in and of itself. Fourth, the teachers in this sample are not a random sample of teacher graduates from EMU, thus we cannot know for certain that all of the relationships uncovered here hold true for the population of EMU teacher education graduates who are currently teaching.

Results

Students

There were 210 mostly Caucasian students in the sample. Of the 210 students 47.1 % were female and 52.9 % were male. The average attendance for these students was 172.99 days ($SD = 4.34$) in a state where 180 days is a normal school year.

Table 1 reports grades and scale scores on the MEAP assessments for the entire sample. It is important to keep in mind that special needs students' grades are typically based on growth rather than achievement, so the grade averages are different when special needs students are excluded from the sample. For the MEAP mathematics test, a score of 521 or higher is required for a "Satisfactory" rating. For the MEAP story selection and information section, the standard of performance is a score of 300 on each section. A "Satisfactory" score *overall* on the reading test can be obtained with a score above 300 on one selection and just below 300 on another selection. Approximately 72% of students received "Satisfactory" scores on the MEAP mathematics test. Around 60% of students received "Satisfactory" scores *overall* on the MEAP reading test.

Table 1

Mean Grades and MEAP Test Scores for the Entire Sample of 4th Graders

	<u>M</u>	<u>SD</u>	<u>n</u>	<u>Scale</u>
Grades in Mathematics	3.13	.65	208	0 - 4.00
Grades in Reading	3.27	.55	208	0 - 4.00
MEAP Mathematics	539.39	30.47	208	463 - 603
MEAP Reading: Story Selection	317.64	23.82	208	240 - 357
MEAP Reading: Information Selection	311.77	21.15	209	237 - 367

When special needs students are removed from the sample, the descriptive statistics don't change very much. These results are found in Table 2.

Table 2

Mean Grades, MEAP Test Scores, and Attendance for Non-Special Needs 4th Graders

	<u>M</u>	<u>SD</u>	<u>n</u>	<u>Scale</u>
Grades in Mathematics	3.22	.60	172	0 - 4.00
Grades in Reading	3.35	.51	172	0 - 4.00
MEAP Mathematics	545.42	27.15	172	463 - 603
MEAP Reading: Story Selection	318	21.11	172	266 - 357
MEAP Reading: Information Selection	315	19.54	173	248 - 367
Average Attendance (Grades 2-4)	172.97	4.45	173	0 - 180

Teachers

Data regarding the experience level, credentialing, and association with EMU of each teacher in the sample are presented in Table 3. The teachers in this sample are, on average, veteran teachers (average experience is from approximately 15 to 19 years). A majority hold master's degrees. And, about half of the second grade teachers have some credential from EMU while about two-thirds of the third and fourth grade teachers have some credential from EMU.

The district, in any given school year, will have 8-11 teachers teaching at each grade level, grades 2-4. Data were collected from 30 teachers. If the student came from outside the district for any given grade level, grades 2-4, then teacher variables were recorded as missing. Consequently, because of student mobility, there is missing teacher data for up to 20% of the students. It should be noted, however, that student mobility is most likely a random factor for this district and there is no reason to believe that the students with missing teacher data are substantively different from the students with complete teacher data in the analyses.

The teachers in this sample are predominantly Caucasian. Their salaries average approximately \$49,000 per year. Additionally, the district is located near several major

universities and is constantly engaged in curriculum development, grants, and the development of pre-service teachers.

Table 3

Teacher Characteristics

	Years Experience (M, SD, n)	Percent BA	Percent MA	Percent with EMU Credential
Grade 2 Teachers	(17.22, 9.33, 12)	44.5	55.5	48.0
Grade 3 Teachers	(15.49, 5.23, 9)	30.1	63.6	63.6
Grade 4 Teachers	(19.39, 9.21, 9)	23.7	76.3	63.6

Note: Percent BA = percent of teachers with BA, Percent MA = percent of teachers with MA, Percent with EMU Credential = percent of teachers who earned at least one of their teaching credentials at EMU

Regression Analyses

Multiple regression analyses were used to answer the research question. Variables were entered by forward entry. A regression model was considered statistically significant when the *p*-value associated with the regression model *F* –statistic was less than or equal to .15. The entry of an individual predictor variable was based on *p*-value equal to or less than .05, using a *t*-test to test each predictor’s regression coefficient.

Regression analyses were conducted using MEAP mathematics scale scores, MEAP story reading scale scores, and MEAP informational reading scale scores as separate criterion variables. For the first set of analyses only teacher variable were included. For the second set of analyses special needs status, gender, and average attendance were added to an equation that included only those teacher variables that were useful in the first set of analyses.

Teaching Variables

The best predictors of MEAP mathematics performance ($F_{(9, 151)} = 2.24, p = .022, R = .343$) were EMU affiliation of the 3rd grade teacher ($\beta = .53, t = 3.32, p = .001$) and the highest credential of the 2nd grade teacher ($\beta = .25, t = 2.26, p = .026$). The model for predicting MEAP story reading performance was not statistically significant. Similarly, the model for predicting MEAP informational reading performance was not statistically significant.

Higher scores on the MEAP mathematics test were associated with students who had EMU-affiliated teachers in 3rd grade and higher credentialed (e.g., BA, MA, etc.) teachers in 2nd grade, on average.

Teaching and Learning Context Variables

For these analyses, special need status, gender, and average attendance (grades 2-4) were added to the equation as student variables. The only teacher variables that were included in these analyses were EMU affiliation of 3rd grade teacher and highest credential of 2nd grade teacher. These two teacher variables were chosen because they proved promising in the first set of analyses.

The best predictors of MEAP mathematics performance ($F_{(5, 155)} = 15.07, p < .001, R = .572$) were EMU affiliation of the 3rd grade teacher ($\beta = .40, t = 4.59, p < .001$) and special needs status of the student ($\beta = -.49, t = -7.32, p < .01$). The best predictors of MEAP story reading performance ($F_{(5, 155)} = 11.33, p < .001, R = .517$) were EMU affiliation of the 3rd grade teacher ($\beta = .22, t = 2.40, p = .018$) and special needs status of the student ($\beta = -.46, t = -6.59, p < .01$). The best predictors of MEAP informational reading performance ($F_{(5, 156)} = 7.60, p < .001, R = .442$) were EMU affiliation of the 3rd grade teacher ($\beta = .22, t = 2.27, p = .024$) and special needs status of the student ($\beta = -.39, t = -5.25, p < .01$).

Higher scores on the MEAP mathematics test were associated with students who had EMU-affiliated teachers in 3rd grade and were not identified with special needs, on average. Higher scores on the MEAP story reading test were associated with students who had EMU-affiliated teachers in 3rd grade and were not identified with special needs, on average. Finally, higher scores on the MEAP informational reading test were associated with students who had EMU-affiliated teachers in 3rd grade and were not identified with special needs, on average.

Discussion

As mentioned in the limitations section of this paper, the results must be interpreted cautiously. This study provides some empirical basis for connecting students' cognitive achievements with teaching variables in general, and their institutional affiliations in particular. It appears that while taking teaching experience, and level of educational attainment (BA, MA, etc.) into account, having credentials from EMU has a positive impact on student achievement in this district, where 48%-66% of elementary grade teachers hold a credential from EMU.

Though there may be concerns that the EMU-affiliated teachers in this sample were given fewer special needs student to teach, the data do not support that contention. A chi-square analysis of the relationship between special needs status of the student and EMU affiliation of the teacher indicate no relationship ($\chi^2_{(1)} = .549, p = .55$). Nor is there a relationship between the special needs status of the student and the classroom to which s/he was assigned ($\chi^2_{(7)} = 2.84, p = .90$). Based on conversations with each principal and

the data presented here, there is no reason to believe that EMU-affiliated teachers were systematically being allowed to teach more able students than non EMU-affiliated teachers.

Research that models achievement at the student level is clearly important. Hierarchical linear models, as well as value-added assessment models, such as those proposed by William Sanders at the University of Tennessee, may prove to be *quite* useful as teacher education institutions seek to meaningfully measure the impact that their graduates have on children in schools.

References

Barton, P., & Coley, R. (1994). *Testing in America's schools* (Policy Information Report). Princeton, NJ: Educational Testing Service.

Bidwell, C., & Kasarda, J. (1980). Conceptualizing and measuring the effects of school and schooling. *American Journal of Education*, 3, 401-429.

Bryk, A., & Raudenbush, S. (1992). *Hierarchical linear models: Applications and data analysis methods*. Sage: London.

Finn, J., & Achilles, C. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis*, 21, 97-110.

Haney, W. (1980). Units and levels of analysis in large-scale evaluation. *New Directions for Methodology of Social and Behavioral Science*, 6, 1-13.

Nye, B., Hedges, L., & Konstantopoulos, S. (1999). The long-term effects of classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis*, 21, 127-142.

Olson, L. (1998). A new accountability player: The local newspaper. Retrieved August 22, 2001. <http://edweek.com/ew/ewstory.cfm?slug=40media.h17&keywords=Newspaper%20Accountability>

Ryan, J. M. (1995). *Current practices and procedures in testing, measurement, and assessment*. Reston, VA: National Association of Secondary School Principals (NASSP).

Wiggins, G. (1992). Creating tests worth taking. *Educational Leadership*, 49(8), 26-33.