Making Sense of the Links: Professional Development, Teacher Practices, and Student Achievement

by Marjorie R. Wallace

Background/Context: Although there is substantial evidence that high-quality professional development can improve teacher practices, less evidence exists for the effects of teacher professional development on intermediate outcomes, such as teacher practices, and their ultimate effects on K–12 student achievement. This work links professional development through teacher practices to examine their separate and combined effects on student achievement.

Research Questions: When teacher characteristics and teacher preparation program are controlled, what are the effects of teacher professional development on (1) teacher practices in mathematics and reading, and (2) subsequent student mathematics and reading achievement?

Population: The study uses students (n = 1,550–6,408) nested within teachers (n = 168–1,029) from six existing databases, two from the 2000 Beginning Teacher Preparation Survey conducted in Connecticut and Tennessee, and four from the National Assessment of Educational Progress (NAEP Mathematics 1996, 2000, and NAEP Reading 1998, 2000).

Research Design: This quantitative study employed a hybrid structural equation model built based on relationships indicated by the literature. Using extant large-scale data sets, the model was first tested using the smallest data set and then confirmed using successively larger state and national data sets.

Conclusions: Professional development has moderate effects on teacher practice and very small but sometimes significant effects on student achievement when the effects of professional development are mediated by teacher practice. In spite of differences in samples, academic subjects, and assessments, the effects of professional development on teacher practice and student achievement persist and are remarkably similar across analyses.

Using multiple data sets and structural equation modeling, this work finds that professional development has moderate effects on teacher practice and very small but sometimes significant effects on student achievement when the effects of professional development are mediated by teacher practice. In spite of differences in samples, academic subjects, and assessments, the effects of professional development on teacher practice and student achievement persist and are remarkably similar across analyses.

The study analyzes six data sets: the Beginning Teacher Preparation Survey (BTPS) conducted separately in two states, Connecticut and Tennessee (Valli, 2000), and consecutive administrations of the National Assessment of Educational Progress (NAEP Mathematics 1996, 2000, and NAEP Reading 1998, 2000). The project asks the following questions at both state and national levels: What are the effects of teacher professional development on (1) teacher practices in mathematics and reading, and (2) subsequent student mathematics and reading achievement? By taking advantage of existing multilevel state and national data collected from 1996 to 2000, the study develops a model that measures the effects of professional development and teacher practices on student achievement in mathematics and reading on a large scale while controlling for teacher preparation program and teacher characteristics.

After a brief look at the research that ties professional development to teaching practices and student achievement in mathematics and reading, the project builds a structural equation model based on the literature. The model is then tested using the smallest data set and confirmed using successively larger state and national data sets. The results are followed by a discussion of findings and implications for the future.

BACKGROUND

This work not only extends the effects of professional development through teacher practices to student achievement, it also controls for complex sources of variation found at the teacher and teacher preparation institution levels. Results from multiple data sets provide an indication of consistency across contexts. Even so, there are potential sources of variation related to context that may cause the relationships among these variables to shift at other times and in other places.

Until the 1990s, education research tended to look at either the relationship between professional development and teacher practice, or the relationship between teacher practice and student achievement for a single academic subject, controlling for a limited number of
covariates. Since that time, more complex studies used multilevel frameworks to capture classroom teacher effects on student achievement or district-level professional development on teachers within schools. Relatively few studies attempted to extend the effects of professional development through teacher practices to student achievement, and results were inconsistent. After a look at study context, there is a short review of recent mathematics and reading research in this area.

STUDY CONTEXT

Effective teaching and learning take place in a world in which policy, national and state governments, institutions of higher education (IHEs), districts, schools, teachers, students, and subject matter all play a part. One key contextual factor contributing noise to the analysis but not fully measured in the data for this study is the complex policy environment surrounding K–12 teaching and learning, which changes from state to state and at the national level. Another key contextual factor not captured in the data is the wide difference of opinion that exists regarding which student outcomes are the most appropriate for judging teacher effectiveness and student learning.

Education in the United States is complicated by a world of competing expectations, demands, and separately administered but intertwined levels of responsibility. A long list of research, both qualitative and quantitative, attests to a resistant and fragmented education system fraught with uncertainty and conflicting goals, in which students, teachers, administrators, and districts work toward different ends (Cusick, 1992; Floden & Buchmann, 1993; Labaree, 1997; Lortie, 1975; Muncey & McQuillan, 1996).

At the local level, principals and teachers may or may not understand or follow state standards and practice guidelines (Spillane, 2005; Youngs & King, 2002). Teachers, especially beginning teachers, often default to the way they were taught when they were students (Lortie). When teachers do seek to improve their practice based on research results, evidence is often lacking on what constitutes high-quality research and what practices actually improve student scores (What Works Clearinghouse, 2002). There are differences of opinion as to which outcomes are appropriate measures of quality teaching and student learning. Many politicians, parents, and stakeholders are insisting on a simple testing and accountability system in which pushing button A will trigger button B to produce the desired result—a single score for which schools, teachers, and student learning can be held accountable (Paige, 2002, 2003, 2004; Traub, 2002). Many researchers and educators hold that large-scale assessments are important tools in measuring the quality of instruction and student learning when those tests are appropriately designed and used for the intended purpose, and they caution against a “one size fits all” mentality (Amrein & Berliner, 2002; Linn, Baker, & Betebenner, 2002). In addition, other researchers and educators feel even more strongly that standardized tests should be a small element of a good learning assessment program but that narrowly focusing on test scores will only stunt the very creativity necessary to success in the global economy (Duckworth, 1987; Zhao, 2006). This study uses existing student test scores as outcomes—both one-time achievement status in NAEP and gains over the course of a year in Connecticut and Tennessee—to determine if such scores can be linked to professional development though teacher practices. Now the overview switches to literature that ties professional development to teacher practice and subsequent student achievement in mathematics and reading.

PROFESSIONAL DEVELOPMENT

Although there is substantial evidence that high-quality professional development can improve teacher practices (D. K. Cohen & Hill, 2001; Desimone, Porter, Garet, Yoon, & Birman, 2002; Kennedy, 1998; National Reading Panel, 2000; Pearson, 1996; Supovitz, 2001), less evidence exists for the effects of teacher professional development on intermediate outcomes, such as teacher practices, and their ultimate effects on K–12 student achievement (Kennedy; Supovitz). In fact, Supovitz’s 2001 review of the literature reveals little evidence that teacher practices are more than moderately related to increases in student achievement on standardized tests, although the effect is positive. This work links professional development through teacher practices to examine their separate and combined effects on student achievement.

TEACHER EFFECTS ON STUDENT ACHIEVEMENT

Several studies assess the size of teacher effects on student mathematics achievement (Goldhaber & Brewer, 1997; Nye, Konstantopoulos, & Hedges, 2005; Rowan, Correnti, & Miller, 2002; Sanders & Rivers, 1996). Results from those studies suggest that at least 21% of the variance in student achievement is related to variation in teacher effectiveness. It should be noted that teacher effectiveness in these studies encompasses, but is not limited to, teacher practices in the classroom. By controlling for teacher practices and teacher preparation program where possible, this study isolates the effects on student achievement attributable to professional development.

Experts disagree about what teacher reading practices are most desirable. Unlike mathematics skills, which are predominantly learned in schools, the basic skills to support reading development are formed at home and in the community (Heath, 1983). The effects of teacher reading practices may be more difficult to interpret because they are confounded by issues of culture and power (Delpit, 1988; Foster & Peele, 1999; Heath; Hirsch, 1996). Most educators agree that a combination of traditional and reform instructional strategies produces the best overall results for most children (Camilli, Vargas, & Yurecko, 2003; Pearson, 1996). The research conducted for this study examines both mathematics and reading achievement to see if there are differences in the effectiveness of professional development and teacher practice on student achievement by subject matter.

With increased national focus on teacher quality and improving results for all children, it is important to understand the unique contributions of professional development to teacher practice and student achievement while taking into consideration other important variables known to contribute to improved teacher practice. Professional development, when effective, is the most promising.
cost-effective tool available to teachers, schools, districts, and states as they seek to improve the quality of teachers (Cohen & Hill, 2001) and subsequent student achievement.

SAMPLES AND VARIABLES

To ensure good answers to the research questions, the study uses six existing databases, two from the Beginning Teacher Preparation Survey (BTPS) conducted in Connecticut and Tennessee and four from NAEP, to answer the same research questions about mathematics and reading. BTPS is a particularly rich data set specifically designed to study the relationships of teacher preparation and professional development to teacher knowledge, practice, and beliefs and to student achievement in Grades 3–8. The NAEP program regularly assesses achievement in mathematics and reading in nationally representative samples of students in Grade 4. Although students are the main focus of the national assessment, teachers fill out questionnaires regarding their education, certification, and teaching practices for all classes that include assessed students. The single most important characteristic shared by the BTPS and NAEP samples is the existence of individual measures of student achievement nested within specific teachers. This allows teachers’ practices in the classroom to provide the all-important link between teacher professional development and student achievement.

Starting with the smallest data set and progressing to the largest makes it possible to confirm results from the first analysis using successively larger samples. The final samples are reported in Table 1.

Table 1. Teacher and Student Samples by Data Set

<table>
<thead>
<tr>
<th>Samples</th>
<th>Mathematics</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher n</td>
<td>168</td>
<td>1550</td>
</tr>
<tr>
<td>Student n</td>
<td>168</td>
<td>227(24)</td>
</tr>
<tr>
<td>Student gain</td>
<td>3.4(7.71)</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Roughly three quarters of BTPS teachers taught in classrooms where students remained for all subjects. Therefore, the Connecticut and Tennessee samples of teachers are analyzed twice, once for student mathematics outcomes and once for student reading. The number of assessed students varies by subject matter. NAEP student and teacher samples are different for each assessment and year and only represent fourth-grade classrooms.

BTPS STATE SAMPLES

BTPS data were collected during the 1999–2000 school year. The states of Connecticut and Tennessee were selected based on their national reputations for innovative K–12 educational practices. BTPS targeted third- through eighth-grade mathematics and reading/language arts teachers in their first 3 years of practice in public schools. It is assumed that beginning teachers are closer to their teacher preparation experiences, so connections to teacher preparation programs are likely stronger than for teachers who have been in the K–12 system for a number of years. Student gain scores (posttest minus pretest) on state mathematics and reading assessments standardized by grade level were tied to specific teachers, and teachers were tied to specific professional development experiences.

Connecticut

BTPS took advantage of an existing sample of beginning teachers (n = 168) selected by the Connecticut State Department of Education (CSDE) for a prior study that looked at student gains in mathematics and reading. Scores existed for approximately 2,500 students nested within 168 teachers. Student data were then aggregated to the teacher level for structural equation modeling analysis.

Tennessee

The Tennessee sample consisted of 442 beginning mathematics and reading teachers from third through eighth grade. Because of confidentiality concerns, Tennessee did not include individual student scores tied to teachers as part of the BTPS data. After aggregating students by teacher using the value-added framework (Sanders & Horn, 1998), student outcome correlations were provided to BTPS researchers who then used them to create a full correlation matrix for all teacher variables.

NATIONAL SAMPLE

At the national level, gain scores do not exist; however, one purpose of NAEP and its complex design is to serve as an effective barometer of change over time. Two consecutive administrations of fourth-grade mathematics (1996 and 2000) and two consecutive administrations of fourth-grade reading assessments (1998 and 2000) provide a measure of change in student achievement and teaching practice over time. No measure of teacher preparation program exists in NAEP.

For all four NAEP data sets, the public school sample with no accommodations was selected because it renders the closest match to the...
student samples collected by the BPTS. Applying the national weight produced national samples of students. The NAEP teacher sample exists as part of the student record so that the data for teachers who have multiple students assessed appear multiple times at the student level. Student data were therefore aggregated to the teacher level. Consequently, the teacher sample is not random. Analyses using these data must be interpreted as random samples of students who have teachers with certain characteristics or who employ certain practices.

MEASURES

Having described the samples used in the study, explanations of variables included in the analysis are in order. All final teacher variables are scales constructed from self-report data. Using the framework originally developed in the BTPS (see Valli, Reckase, Raths, & Rennert-Ariev, 2001), similar scale variables were created with NAEP data (Wallace, 2005). Scale variables fall into three main conceptual categories: professional development, teacher practice, and student achievement. Professional development is believed to affect teacher practice, which in turn affects student achievement.

Professional development

Professional development allows a teacher to hone current skills and knowledge, as well as to keep abreast of new knowledge, theories, and methods (Borko & Putnam, 1996). Professional development also has the ongoing potential to support and improve the in-service teaching force. In this study, professional development is a combination of mentoring experiences that teachers received, and professional development sessions that they attended since completing formal teacher preparation. Scales for professional development impact, professional development frequency, and mentoring worth were developed as part of the original BTPS study. Of the three, only professional development frequency is available in NAEP Mathematics and Reading across study years. Professional development frequency indicates the amount of time spent in the activities. Professional development impact includes ratings of how helpful teachers found various activities commonly covered in professional development, such as using a wider range of strategies to help students learn or managing the classroom more effectively. Mentoring worth measures how teachers feel about their mentoring experiences and ranges from not at all valuable to very valuable.

Teacher practices

Teacher practices in mathematics and reading are some of the most important influences on student learning. It is in the classroom that pedagogical content knowledge (Shulman, 1987), the combination of subject matter knowledge and pedagogical skills, is synthesized into a cohesive lesson that students can understand. Pedagogical content knowledge also enables the teacher to elicit and/or circumvent common misconceptions surrounding the subject matter (Borko & Putnam, 1996).

Mathematics. The 2000 National Council of Teachers of Mathematics [NCTM] standards (National Council of Teachers of Mathematics, 2000) recommend a combination of strategies, both conceptual and computational, as important to student mathematics achievement. BTPS and NAEP scales for teacher mathematics practices covered three important types of activities that include both conceptual and computational elements: mathematics group and project activities, mathematics drill and lecture activities, and mathematics sense-making activities.

Mathematics group and project activities combine items such as working in small groups to solve math problems and working on mathematical investigations that span several days. Items like lecturing or presenting information and drilling students on computational skills are included in mathematics drill and lecture activities. The mathematics sense-making activities scale reflects the frequency of conceptual activities that teachers employ in the classroom to help students make sense of mathematics (i.e., having students make conjectures, leading discussions on ways students solve particular problems, applying math to real-world problems, modeling different learning strategies, and so on). Teachers’ emphasis on group and project assessment was substituted in the NAEP Mathematics 1996 analysis.

Reading. Reading practices measure the frequency that teachers use reform activities, traditional activities, computer activities, and literacy materials in their classrooms. Reform activities include comprehension activities, metacognitive routines for independent reading, oral and written responses to literature, and so on. Traditional activities encompass activities such as oral and silent reading and reading vocabulary. Most teachers use both types of activities in the classroom. The scale for literacy materials is available in BTPS and shows the range of textbook use from none to a great amount. Frequency of computer activities is an additional variable available in NAEP.

Student achievement

Data sets in this study have one of two outcome measures for student mathematics and reading: gain scores in BTPS and multiple plausible values in NAEP. Although student gain scores have been criticized because of the many influences besides teaching that can affect student scores over the course of a single year (Linn et al., 2002), gains over 3 years have been successfully employed as a barometer of growth in Tennessee (Sanders & Horn, 1998). Each student essentially serves as his own control group, eliminating many family and societal factors external to the classroom. The CSDE supplied two scores on the Connecticut Mastery Test per student, one from the beginning and one from the end of the 1999–2000 school year for all teachers involved in the BTPS. Gains scores were computed by subtracting pretest results from posttest results and standardized by grade. When using gain scores, it is important to note that the effects of independent variables on the outcome are likely to be underestimated (Rowan et al., 2002).
The outcome variable in NAEP is a measure of achievement status, a one-time score. NAEP is unusual in that no student completes the entire assessment. Using a complex system of weighting and matrix sampling, five plausible values representing a range of possible scores are created for each student based on like student patterns found in the data. For the structural equation modeling analyses, the five plausible values (MMPCM1-5 for mathematics or RRPCM1-5 for reading) are included as indicator variables for the outcome variable, student achievement.

Final data preparation

Because the variables for Tennessee were only available in the form of a correlation matrix, correlation matrices were created for the remaining data sets using the variables just presented (Wallace, 2005). Special weighted correlation matrices were constructed for NAEP to ensure appropriate standard errors (J. Cohen, 2002).

METHODS

To determine if professional development and teacher practice were adequately represented by the selected scale variables, confirmatory factor analysis was necessary. Estimation of professional development effects on teacher practices and subsequent student achievement required a path model. Structural equation modeling is able to perform both analyses simultaneously. Figure 1 presents the generic structural equation used for all analyses.

Figure 1. Generic Structural Equation Model

The relationships among professional development, teacher practice, and student achievement are shown in bold, and controls for teacher preparation program and teacher characteristics are shown in light gray. Only teacher practice is hypothesized to be directly related to student achievement/gains. Professional development, as well as the control variables, must go through classroom practice to affect student achievement. Ovals represent key conceptual variables, and rectangles represent scale variables actually measured in the data. When the results are standardized, the coefficients are called effects with a mean of 0 and a standard error of 1. Positive effects mean greater than average effects, and negative effects mean less than average effects.

One criterion for judging effect size in education research is its potential value for informing or benefiting education practice (Camilli et al., 2003; McCartney & Dearing, 2002; McCartney & Rosenthal, 2000). Given the complexities of the educational system and the many intervening circumstances that influence teachers and students inside and outside the classroom, effect sizes are expected to be small. BTPS researchers anticipated effects sizes of no more than .05 over the course of 1 year, and they considered this a potentially significant finding (Valli, Reckase, & Raths, 2003). Generally, absolute values less than .10 may indicate small effects for a standardized path coefficient; absolute values around .30 are medium; and absolute values of .50 or more are considered large (Kline, 1998).

RESULTS

To answer questions regarding the effects of professional development and teacher practice on student achievement outcomes, the analysis began with structural equation models for mathematics and reading in Connecticut, the smallest data set. These were followed by the analyses in Tennessee and NAEP. In all, there were eight structural equation models across six data sets. The first part of this section interprets $r^2$ results, which reveal the relative importance of the individual scale variables to key hypothetical concepts; the second part provides interpretation of fit statistics; and the third part outlines effect sizes for professional development and teacher practices, followed by a description of study limitations.

The shared variance ($r^2$) between the scale variables and a conceptual variable, such as professional development, determines the relative
importance of the individual characteristic. Table 2 compares $r^2$ for scale variables across the analyses. Generally, $r^2 > .35$ is considered large, $r^2 > .15$ is moderate, and $r^2 > .02$ is small when interpreting indicator variables (J. Cohen, 1988).

### Table 2. Comparison of $r^2$ of Indicator Variables for Professional Development, Teacher Practice, and Student Achievement in Mathematics and Reading Across Data Sets.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Mathematics</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional Development</td>
<td>$r^2 = .372$</td>
<td></td>
</tr>
<tr>
<td>Mentoring worth</td>
<td>$r^2 = .303$</td>
<td></td>
</tr>
<tr>
<td>Professional development frequency</td>
<td>$r^2 = .776$</td>
<td></td>
</tr>
<tr>
<td>Mathematics sense-making activities</td>
<td>$r^2 = .944$</td>
<td></td>
</tr>
<tr>
<td>Reform reading activities</td>
<td>$r^2 = .611$</td>
<td></td>
</tr>
</tbody>
</table>

An example of how to interpret Table 2 follows. Professional development in Connecticut is described by variables for mentoring worth ($r^2 = .372$), professional development frequency ($r^2 = .303$), and professional development impact ($r^2 = .776$). A large amount (over 75%) of the variance is shared between professional development and the variable for professional development impact, which means that it is an effective indicator for this hypothetical concept. Mentoring worth and professional development frequency share moderate to large amounts of variance with the construct as well.

Looking across both mathematics and reading results, Connecticut and Tennessee analyses indicate that professional development impact ($r^2 = .739–.862$) is a larger and more consistent indicator of professional development than professional development frequency ($r^2 = .094–.303$). So is mentoring worth ($r^2 = .268–.372$). In NAEP analyses, professional development frequency is analyzed as a single predictor variable.

Variables sharing variance with teacher practice in both mathematics and reading ranged from $r^2 = .006$ to .944 but were predominantly in the medium to large range. Mathematics sense-making activities and reform reading activities shared the most variance with teacher practices for mathematics and reading, respectively, and were therefore good indicators. Mathematics drill and lecture activities ($r^2 = .006–.114$) and use of literacy materials ($r^2 = .006–.061$) were weak indicators for teacher practice in these data.

Gains were measured as continuous outcome variables in Connecticut and Tennessee. Achievement status in NAEP has five imputed plausible values (essentially five possible scores had the student been given all items on the test) that serve as the indicator variables ($r^2 = .900–.966$).

### Table 3. Standardized Total Effects for Professional Development and Teacher Practice on Mathematics Achievement

Results are standardized (mean = 0, SD = 1) to aid interpretation across data sets. The magnitude of an effect is its absolute value; therefore, positive effects represent above-average effects, and negative effects represent below-average effects. Table 3 may be read as follows. The first column under Professional Development includes the effects of professional development on teacher practice ($^*d = .166$) and the effects of professional development on student mathematics gains moderated by teacher practice for Connecticut ($^*d =$
The second section (Teacher Practice) shows the factor loadings and the effects of teacher practice on average student gains/achievement across all data sets. Fit indices to the right exhibit good fit and parsimony across models and data sets. As expected, the fit indicators become stronger with increasing sample size, indicating that the model is a reasonable interpretation of the data.

The structural equation model for Connecticut mathematics (Figure 2) displays the effects from Table 3 positioned near their corresponding paths. Similar models were analyzed for each data set and subject (see Wallace, 2005). Variations in variables across data sets required some minor modifications to variables represented across models.

**Figure 2. Connecticut Structural Equation Modeling Mathematics Model With Standardized Effects**

Effect sizes for reading are shown in Table 4. The table may be interpreted in the same way as Table 3.

<table>
<thead>
<tr>
<th>Table 4. Standardized Total Effects for Professional Development and Teacher Practices on Reading Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teacher x</strong></td>
</tr>
<tr>
<td><strong>Professional Development</strong></td>
</tr>
<tr>
<td>Monitoring with professional development</td>
</tr>
<tr>
<td>Professional development frequency</td>
</tr>
<tr>
<td>Professional development impact</td>
</tr>
<tr>
<td><strong>Teacher Practice</strong></td>
</tr>
<tr>
<td>Reading Activities</td>
</tr>
<tr>
<td>Traditional Activities</td>
</tr>
<tr>
<td>Computer Activities</td>
</tr>
<tr>
<td>* Literacy materials</td>
</tr>
<tr>
<td><strong>Reading Achievement Outcomes</strong></td>
</tr>
</tbody>
</table>

*Significance: ***p < .001; **p < .01; *p < .05.

The generic structural equation modeling model presented in Figure 1 served well across all data sets. Each latent variable had a series of descriptor variables that formed separate measurement models. The measurement models successfully defined the latent variables for professional development, teacher practice, and student achievement. When the measurement models were combined with the structural model, the final hybrid model exhibited reasonably good fit (NFI > .984, p < .001, root mean square error of approximation = .05–.116) and worked in similar ways across data sets. Variations in effects that occurred appeared related to subject matter and state versus national data set. The most important findings from the tables are organized by research question and presented next.

**ASSESSING OVERALL MODEL FIT**

The generic structural equation modeling model presented in Figure 1 served well across all data sets. Each latent variable had a series of descriptor variables that formed separate measurement models. The measurement models successfully defined the latent variables for professional development, teacher practice, and student achievement. When the measurement models were combined with the structural model, the final hybrid model exhibited reasonably good fit (NFI > .984, p < .001, root mean square error of approximation = .05–.116) and worked in similar ways across data sets. Variations in effects that occurred appeared related to subject matter and state versus national data set. The most important findings from the tables are organized by research question and presented next.

**EFFECTS OF PROFESSIONAL DEVELOPMENT**
Across the mathematics analyses, a one-standard-deviation increase in professional development translates into moderate increases to average frequency of teacher mathematics practices occurring in the classroom. The effects of professional development on teacher mathematics practice, controlling for teacher preparation program and teacher characteristics at the state level, are moderate (d* = –.12–.17). In NAEP, in which teacher characteristics are controlled but it is not possible to control for teacher preparation program, the effects of a one-standard-deviation increase for professional development (about 1.2 hours per year) on teacher practice are still in the moderate range (d* = –.24–.34).

In reading, the effects of professional development on teacher practice, controlling for teacher preparation program (Connecticut and Tennessee only) and teacher characteristics, are not as large or consistent as in mathematics (d* = –.03–.16). A one-standard-deviation increase in professional development at the state level is related to small increases in average frequency of teacher reading practices used in the classroom (d* = –.12–.16). At the national level, frequency of professional development (1 SD = about 1.1 hours per year) also has small effects on the frequency of reading practices used in the classroom. In NAEP 1998, however, a one-standard-deviation increase in professional development frequency is related to a decrease in average frequency of reading practices in the classroom (d* = –.07). There is a link between professional development and teacher reading practice, but the results indicate that it is as likely to support and increase practices as lead to a reduction in the use of those practices.

In general, professional development has very small but occasionally significant effects on average student achievement in mathematics and reading when mediated by teacher practice. A one-standard-deviation increase in professional development controlling for teacher preparation program and/or teacher characteristics is consistently related to very small increases to average mathematics gains and achievement status at state and national levels (d* = 0.01–0.03). In Connecticut, mathematics gains increased by about .07 points, and in Tennessee, gains increased by about .15 points. In NAEP 2000 mathematics, for instance, this means that an additional 1.2 hours of professional development per year is related to increases in average scores of .62 points. In NAEP 1998 reading, 1.1 hours of professional development is related to increases in average scores of .24 points.

Professional development effects on reading gains and achievement mediated by teacher practice are very close to zero and range from d = –.01 to .01, with one exception: Connecticut. The largest effect for professional development on student achievement occurs in reading in Connecticut (d = –.05). This means that a one-standard-deviation increase in teacher professional development is related to a .39-point increase in student reading gains, representing an increase of about 11% over the course of 1 year. In Connecticut, a state with a reputation for standards alignment and nationally high reading achievement, increases in the frequency and quality of professional development lead to increases in reading practices, which translate into significant reading gains.

EFFECTS OF TEACHER PRACTICE

This study did not set out to examine the effects of teacher practice on student achievement, yet the relationship has important implications. The direct effects of teacher practice on student achievement, controlling for teacher preparation program, teacher characteristics, and professional development, are larger than the indirect effects of professional development on student achievement. Teacher practice has small to moderate direct effects on average student mathematics achievement across all data (d* = .042–.170), which translates into an increase in average gains of 8.5% in Connecticut and an increase in average achievement status of 1% in NAEP.

Again, there are inconsistencies in the results for reading. A one-standard-deviation increase in teacher reading practice controlling for teacher preparation program, teacher characteristics, and professional development in Connecticut is related to an increase in average elementary reading achievement of 65%. At the other extreme, teacher reading practice in NAEP 1998 also has a moderate direct effect (d* = –.125), which translates into a 2% decrease in average student reading achievement.

It is worth noting that 68%–85% of the variance in teacher mathematics practice is yet to be explained by the structural equation models in this study. Fifty percent to seventy-five percent of the variation in teacher reading practices is left to be explained. Atypical, the NAEP 1998 model leaves only 18% of the variation in reading practices to be explained. In this study, unexplained variance represents unmeasured variables that may be related to teachers’ personal backgrounds, individual styles, peer influences, school climate, and other factors not included in the study but that clearly appear to have great influence on the frequency of practices that teachers use.

LIMITATIONS OF DATA AND METHODS

Factors that may affect interpretations of the analysis include large variations in sample size and nonrandom assignment of teachers and/or students. Sample size ranged from 168 teachers and 1,550 students in Connecticut to over 1,000 teachers and 6,000 students in NAEP. It is well-known that larger sample sizes have more stability and accuracy than smaller ones. To improve interpretation across differing sample sizes, this analysis proceeded with the smallest data set serving as the base analysis, with results confirmed in successively larger data sets.

Students and teachers are rarely randomly assigned to classrooms. In the data used for this study, the Connecticut teacher sample was randomly selected by the state. The Tennessee sample was a convenience sample of those beginning teachers who responded to the survey. In NAEP, students were randomly selected, but teachers from all experience levels were not. If multiple sources of evidence confirm analysis results, it is likely that these threats to validity are minimized, but it is still questionable to assert causal inferences.
A methodological choice that may have affected interpretation of results is the use of correlation matrices as input for structural equation modeling. Although covariance matrices are the considered a better choice for data format, the inherited correlation format of the Tennessee data dictated input format for the other analyses.

CONCLUSIONS AND DISCUSSION

What effect does teacher professional development have on teacher practice? Does teacher professional development ultimately translate into increased student achievement? Is it the same for mathematics and reading? The answers to these questions hinge on a reliable model that can tie professional development through teacher practice to student achievement while taking into consideration the complexity of the American education system by accounting for factors such as teacher preparation program and teacher characteristics. Using the results from eight structural equation analyses representing two state and four national data sets, this research (1) demonstrates that a hybrid structural equation model may be used to link professional development to teacher practice and estimate resulting effects on student achievement across multiple data sets, subject matters, and assessment levels, (2) finds that the effects of professional development on mathematics and reading practice are generally moderate, and the effects of professional development on student achievement mediated by teacher practice are very small but sometimes significant, (3) finds that substantial teacher practice effects on student achievement are still unexplained when professional development, teacher preparation program, and teacher characteristics are controlled, and (4) concludes that the education system is better aligned to promote mathematics achievement than it is to promote reading achievement.

THE MODEL

Newly implemented No Child Left Behind (NCLB) accountability measures (No Child Left Behind Act, 2001) make it increasingly important to develop complex models that acknowledge the complicated nature of education while yielding useful answers to important questions regarding the quality of teaching and learning for all students. This study developed a complex model that used structural equation modeling to analyze the effects of teacher professional development on teacher practice and subsequent student achievement in two subjects across six data sets representing both state and national samples. The model also controlled for teacher preparation program (state level only) and teacher characteristics.

Results suggest that teacher professional development, teacher practice, and student mathematics/reading achievement are connected in similar ways at the state and national levels. Model stability across different conditions while maintaining goodness of fit statistics demonstrates stable connections among variables, which in turn allows for the comparison of effect sizes across analyses. Although the effects of professional development and teacher practice on student achievement can be measured, some effects are strong and consistent, and others are not. Other conclusions springing from the research are summarized and discussed below.

Professional development effects can be linked to teacher practice and subsequent student mathematics and reading achievement. Reviews of literature pointed to the lack of research on the links from teacher professional development through teacher practice to student achievement (Supovitz, 2001; Wayne & Youngs, 2003; Wilson, Floden, & Ferrini-Mundy, 2001). This study demonstrates not only that it is possible to include these elements in a single model, but also that the effects of professional development on teacher practice are small to moderate, with very small but occasionally significant indirect effects on student outcomes.

Prior studies that measured teacher effects on student mathematics achievement indicated that at least 21% of the variation in student outcomes was related to variation in teacher effectiveness (Goldhaber & Brewer, 1997; Nye et al., 2005; Rowan et al., 2002; Sanders & Rivers, 1996). The research presented here isolated the effects of professional development by controlling for teacher preparation program and teacher characteristics when possible, so the remaining effects on student outcomes were expected to be small. Small increases in average mathematics gains/achievement do occur across analyses as a result of a one-standard-deviation increase in professional development. The biggest surprise was in reading where a one-standard-deviation increase in professional development raised average student gains (11% in Connecticut), whereas other outcomes remain nearly unchanged at the average.

Understanding the links among variables is critical to addressing how professional development works to produce effective teacher practices that lead to increased K–12 student achievement in mathematics and reading. Although this study focused on the state of the links in education prior to NCLB, mandated student databases will soon be available in every state, offering important opportunities to study these effects further.

Most of the variance in teacher mathematics and reading practice is not accounted for by professional development. Structural equation modeling not only allows estimation of the effects of key conceptual variables, but it also indicates how much of the variance is not explained by those variables. Sixty-eight percent to eighty-five percent of the variance in teacher mathematics practice is yet to be explained by these analyses. Approximately 50%–75% of the variation in teacher reading practices is left to be explained. The unexplained variance represents unmeasured teacher factors that affect classroom practice and may be related to personal backgrounds, individual styles, peer influences, and so on. Future research needs to address the missing pieces to this puzzle.

Subject matter context makes a difference. Although it is not new that different subjects require different teaching methods (Stodolsky, 1998), it is new that subject matter context (mathematics or reading) plays a role in explaining how the various elements in education are connected to student achievement across data sets. Variables are linked to mathematics achievement and less well linked to reading achievement.
Mathematics exhibits that the effects for professional development translate into increased effects on teacher practices and student achievement across all data sets. A one-standard-deviation increase in the predictor variables increases student achievement.

With the exception of Connecticut, effects for reading are inconsistent across data sets, and effects on student achievement are near zero. This implies that professional development is not aligned to support teacher reading practices and reading achievement in the same way that it is aligned to support mathematics achievement. This supports recent findings (Harris & Sass, 2007) noting that professional development has significant effects on student mathematics achievement in Florida but that the effect does not carry over into reading achievement across grades.

Consistent effects for reading in Connecticut suggest that the time and effort that Connecticut spent aligning its K–12 student standards with teacher preparation programs and teacher professional development during 1990s was time well spent (Wilson, Darling-Hammond, & Berry, 2000). A one-standard-deviation increase in teacher preparation program or professional development translates into increases in both the studied teacher reading practices and gains. Lessons from Connecticut will be valuable as other states move toward comprehensive reading policy to reach national reading achievement requirements.

**IMPLICATIONS FOR THE PRESENT AND FUTURE**

As the United States looks to increase K–12 student achievement for all students, knowing that it is possible to measure effects spanning teacher professional development to student achievement in a single model sensitive to subject matter context will prove useful. State and national databases collected since 2000 promise important insights on strengthened relationships and improved outcomes throughout the system. Additionally, if the education community wishes to study the effects of a teacher’s university preparation and professional development on K–12 student achievement, it will be necessary to collect and link state student achievement data with teacher data and important characteristics of their respective preparation programs and professional development experiences.

More research needs to be done with data that span government policy to student achievement. This means that appropriate databases and/or ways to bridge data sets need to be developed. Stronger connections need to be made throughout the education system, especially in reading, so that all elements of the system are working in concert. Models like the one presented here will be useful in exploring the validity of the logic models behind public policy.

**Acknowledgements**

This research was supported by a grant from the American Educational Research Association, which receives funds for its AERA Grants Program from the National Center for Education Statistics and the Institute for Education Sciences (U.S. Department of Education) and the National Science Foundation under NSF Grant No. REC-9980573. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

I want to thank Bob Floden for his unflagging support from predissertation to print. In addition, thanks to Gary Natriello, Robert Calfee, and Mark Hamilton for encouragement and comments on early drafts. Finally, I would like to thank all the anonymous TCR reviewers for insightful suggestions throughout this process.

**Note**

1. To obtain questionnaires before 2003, please contact the NCES Assessment Division at 202-502-7420 or 202-502-7400, or specify the questionnaires you need by logging onto http://nces.ed.gov/nationsreportcard/contactus.asp.

**References**


U.S. Department of Education.


Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the *Prospects* study of elementary schools. *Teachers College Record, 104*, 1525–1567.


Cite This Article as: Teachers College Record Volume 111 Number 2, 2009, p. -