Chapter 5: Report of the Task Group on Teachers and Teacher Education

Deborah Loewenberg Ball, Chair
James Simons
Hung-Hsi Wu
Raymond Simon, Ex Officio
Grover J. “Russ” Whitehurst, Ex Officio
Jim Yun, U.S. Department of Education Staff

CONTENTS

List of Tables 5-v
List of Abbreviations 5-vii

Executive Summary 5-ix
I. Introduction 5-1
II. Methodology 5-2
  A. Categories of Studies 5-3
    1. High-Quality Studies 5-3
    2. Moderate-Quality Studies 5-4
    3. Lesser-Quality Studies 5-4
  B. Procedures 5-4
    1. Literature Search and Study Inclusion 5-4
    2. Effect Size Calculations 5-5
III. Findings 5-6
  A. Teachers' Knowledge of Mathematics 5-6
    1. Introduction 5-6
    2. Teacher Certification as a Measure of Mathematical Content Knowledge 5-7
      a. Overall Findings 5-7
      b. Strength of the Findings 5-10
      c. Magnitude of the Findings 5-10
    3. Content Course Work and Degrees as Measures of Mathematical Content Knowledge 5-11
      a. Overall Findings 5-12
      b. Strength of the Findings 5-13
      c. Magnitude of the Findings 5-13
    4. Test Scores and Ad Hoc Measures as Measures of Mathematical Content Knowledge 5-15
      a. Overall Findings 5-16
      b. Strength of the Findings 5-17
      c. Magnitude of the Findings 5-17
    5. The Mathematical Content and Nature of Teacher Licensure Exams 5-19
      a. Implications from the Empirical Evidence 5-21
    6. Recommendations Based on the Mathematical and Logical Analysis of the Demands of Teaching Mathematics 5-22
  B. Teachers' Education: Teacher Preparation and Alternative Pathways to Teaching, Professional Development, and Induction 5-23
    1. Preservice Teacher Preparation 5-24
    2. Alternative Pathways to Teaching 5-27
    3. Induction Programs 5-33
    4. Professional Development Programs 5-35
    5. Conclusions 5-40
    6. Recommendations 5-41

C. Teacher Incentives 5-41
1. Utilizing Labor Market Incentives for Good Teaching................................. 5-42
a. Skills-Based Pay... 5-42
b. Location Pay... 5-43
c. Performance Pay... 5-43
2. Recommendations 5-47
3. Cautions 5-47
D. Elementary Mathematics Specialist Teachers 5-51
1. What Models Exist for Elementary Mathematics Specialist Teachers and Their Preparation?.. 5-52
a. The Lead Teacher or Math Coach Model.. 5-54
b. The Specialized-Teacher Model.. 5-55
c. The Pull-Out Model... 5-55
2. What Evidence Exists for the Effectiveness of Elementary Mathematics Specialist Teachers with Respect to Student Achievement? 5-55
a. Costs Associated with Mathematics Specialists..................................... 5-56
b. Mathematics Specialists Internationally 5-58
3. Conclusions.. 5-58

BIBLIOGRAPHY 5-59

Tables

Table 1: Quality Characteristics of Models Examining Impact of Teacher Certification on Student Achievement in Mathematics, by Study and by Overall Quality 5-9
Table 2: Reported Impacts of Models Examining the Effect of Teacher Certification on Student Achievement in Mathematics for the High-Quality Studies 5-11
Table 3: Quality Characteristics of Models Looking at Impact of Teacher Math Course Work and Degrees on Student Achievement in Mathematics, by Study and by Overall Study Quality 5-14
Table 4: Reported Impacts of Models Examining the Effect of Teacher Mathematics Course Work and Degrees on Student Achievement in Mathematics, by Study and by Overall Quality 5-15
Table 5: Quality Characteristics of Models Looking at Impact of Teacher Test Scores and Other Ad Hoc Measures on Student Achievement in Mathematics, by Study and by Overall Study Quality 5-18
Table 6: Reported Impacts of Models Examining the Effect of Teacher Test Scores and Other Ad Hoc Measures on Student Achievement in Mathematics, by Study 5-19
Table 7: Quality Characteristics of Models Examining Impact of Teacher Preparation Programs on Student Achievement in Mathematics or Teachers' Mathematics Content Knowledge 5-25
Table 8: Reported Impacts From Models Examining the Effect of Teacher Preparation Programs on Student Achievement in Mathematics or Teachers' Mathematics Content Knowledge 5-26
Table 9: Quality Characteristics of Models Examining Impact of Alternative Pathways on Student Achievement in Mathematics 5-29
Table 10: Reported Impacts of Studies Examining the Effect of Teachers from Alternative Paths on Student Achievement in Mathematics........ 5-30
Table 11: Quality Characteristics of Studies Examining Impact of Teacher Professional Development Programs on Student Achievement in Mathematics........ 5-36
Table 12: Reported Impacts of Studies Examining the Effect of Teacher Professional Development Programs on Student Achievement in Mathematics........ 5-37
Table 13: Quality Characteristics of Studies Examining Impact of Teacher Pay for Performance on Student Achievement in Mathematics 5-45
Table 14: Reported Impacts of Models Examining the Effect of Teacher Pay for Performance on Student Achievement in Mathematics 5-48

Abbreviations

AIMS PK–16	The Alliance for Improvement of Mathematics Skills—PreKindergarten Through Grade 16
ANOVA	Analysis of Variance
BA	Bachelor of Arts
BNSE	Belize National Selection Exam
BSAP	Basic Skills Assessment Program
CCTDM	Classroom Center Teacher Developmental Mathematics
CKT-M	Content Knowledge for Teaching Mathematics
CSAP	Colorado Student Assessment Program
CTBS	Comprehensive Test of Basic Skills
ETS	Education Testing Service
FCAT	Florida Comprehensive Achievement Test
GLS	Generalized Least Squares
GPA	Grade Point Average
GSCE	Graduate School College Education
HLM	Hierarchical Linear Modeling
HS	High School
ICC	Intra-Class Correlation
IMA	Integrated Mathematics Assessment
IP	Incentive Plan
ISD	Independent School District
ITBS	Iowa Test of Basic Skills
KCPE	Kentucky Council on Postsecondary Education
LSAY	Longitudinal Study of American Youth
MA	Master of Arts
MAT	Master of Arts in Teaching
MMI	Maysville Mathematics Initiative
NAEP	National Assessment of Educational Progress
NBPTS	National Board for Professional Teaching Standards
NCATE	National Council for the Accreditation of Teacher Education
NCE	Norm Curve Equivalent
NELS	National Educational Longitudinal Survey
NWEA	Northwest Evaluation Association Study
OLS	Ordinary Least Squares
PFP	Pay for Performance
PK–16	Pre-School Through Grade 16
PPST	Pre-Professional Skills Tests
PS	Propensity Score
RCT	Randomized Controlled Trials
RDD	Regression Discontinuity Design
SAT	Scholastic Aptitude Test
SAT-9	Stanford Achievement Test, Ninth Edition
SC	South Carolina

SSCI	Social Sciences Citation Index
STAR	Student/Teacher Achievement Ratio
SUPP	Supplemental
TAAS	Texas Assessment of Academic Skills
TFA	Teach for America
TIMSS	Trends in Mathematics and Science Study
TRAD	Traditional
WCET	Weighted Common Examinations Total
WWC	What Works Clearinghouse

Executive Summary

Introduction

Teachers are crucial to students' opportunities to learn mathematics, and substantial differences in the mathematics achievement of students are attributable to differences among teachers. Therefore, the National Mathematics Advisory Panel (Panel) was charged with making recommendations, based on the best available scientific evidence, on "the training, selection, placement, and professional development of teachers of mathematics in order to enhance students' learning of mathematics," according to presidential Executive Order 13398. To address this charge, the Panel established a Teachers and Teacher Education Task Group (Task Group), which combed the research on the relationship between teachers' own knowledge and students' achievement, and how effective teachers can best be recruited, prepared, supported, and rewarded.

The four questions that structured the work of the Task Group are:

1) What is the relationship between the depth and quality of teachers' mathematical knowledge and students' mathematics achievement?

2) What is known about programs that help teachers develop the necessary mathematical knowledge for teaching? Which of these programs have been shown to affect instructional practice and student achievement?

3) What types of recruitment and retention strategies are used to attract and retain highly effective teachers of mathematics? How effective are they?

4) What evidence exists for the effectiveness of elementary mathematics specialist teachers with respect to student achievement? What models exist for elementary mathematics specialists and their preparation?

The research base that addresses these questions was found to be uneven. Therefore, in addition to making recommendations based on the best available evidence, the Task Group makes recommendations about the research needed in order to improve practice and policy with respect to teachers.

Methodology

The Task Group identified available scientific evidence, published in peer-reviewed journals and national reports, and categorized the quality of the research studies related to each of the four research questions. Studies were categorized as high quality, moderate quality, or lesser quality based on the appropriateness of the design in answering the specific research question. High-quality studies were those that employed a randomized control trial design or those that addressed the weaknesses of a correlational design through the use of large samples, control variables, multiple specifications, etc. Standardized regression coefficients or standardized mean differences were calculated as appropriate. High-quality studies served as the primary basis for the Task Group's recommendations, although the available evidence varied for each research question. Because of the paucity of rigorous empirical research to answer the Task Group's fourth question, the Task Group provides a description of the programs and models that exist in the United States, some distinctions among the different models, and commentary on the costs and benefits of those different models.

Teachers' Knowledge of Mathematics

Common sense dictates that teachers must know the mathematical content they teach, but defining a precise body of mathematical knowledge that would effectively serve teachers and would guide teacher education, professional development, and policy has proved challenging. Therefore, the Task Group considered the empirical evidence on how the level and scope of teachers' own mathematical knowledge affects the learning of their students. Across the rigorous research studies reviewed on the relationship between teachers' mathematical knowledge and their students' achievement, teacher content knowledge was measured in three different ways: teacher certification, mathematics course work, and tests of teachers' mathematical knowledge. Across all studies, the general results are mixed but overall do confirm the importance of teachers' content knowledge. However, because most studies have relied on proxies for teachers' mathematical knowledge (such as teacher certification or courses taken), existing research does not reveal the specific mathematical knowledge and instructional skill needed for effective teaching, especially at the elementary and middle school level. Direct assessments of teachers' actual mathematical knowledge provide the strongest indication of a relation between teachers' content knowledge and their students' achievement.

Teacher Certification as a Measure of Mathematical Content Knowledge

Across studies that used teacher certification or teacher certification in mathematics as a proxy for teachers' mathematical content knowledge, findings were mixed about the impact of teacher certification on student achievement in mathematics. Research in this area has not provided consistent or convincing evidence that students of teachers who are certified in mathematics gain more than those whose teachers are not.

Content Course Work and Degrees as Measures of Mathematical Content Knowledge

Mathematics course work and field-specific degrees are a second common group of proxies for teachers' mathematical content knowledge; both measures focus on teachers' completion of college-level mathematics study and are often jointly considered within the same analysis. In general, although results are mixed, there appears to be some positive relationship between teacher content course work and degrees, and student math achievement at the high school level. However, the existing research does not show evidence of such a relationship below ninth grade.

Test Scores and Ad Hoc Measures as Measures of Mathematical Content Knowledge

More proximal measures of teachers' mathematical content knowledge included tests of teachers' content knowledge. Such measures allow closer examination of the effect that mathematical knowledge has on student achievement. Although there was variation among the set of studies that used teacher test scores as measures of teacher content knowledge, overall these studies signaled a positive effect of mathematical content knowledge on student achievement. It should be noted that the studies this Task Group found were focused at the elementary level, making comparisons with other findings difficult.

The Mathematical Content and Nature of Teacher Licensure Exams

Although recent research treating teacher licensure as a proxy for teachers' mathematical content knowledge has not consistently or convincingly shown that students of teachers who are licensed in mathematics gain more academically than those whose teachers are not, teacher licensure exams play an important role in determining the math teachers available for employment in schools. Recognizing the importance of teacher licensure exams, the Task Group sought access to these exams together with data on teachers' performance on exam items. Due to issues of confidentiality, however, the Task Group was not able to gather information sufficient to assess the mathematical quality of these exams.

Conclusions and Recommendations Regarding Teachers' Knowledge of Mathematics

Research on the relationship between teachers' mathematical knowledge and students' achievement supports the importance of teachers' content knowledge in student learning. However, because most studies have relied on proxies for teachers' mathematical knowledge (such as teacher certification or courses taken), existing research does not reveal the specific mathematical knowledge and instructional skills needed for effective teaching, especially at the elementary and middle school level.

Direct assessments of teachers' actual mathematical knowledge provide the strongest indication of a relation between teachers' content knowledge and their students' achievement. More precise measures are needed to specify in greater detail the relationship among elementary and middle-school teachers' mathematical knowledge, their instructional skill, and students' learning.

Teachers' Education: Teacher Preparation, Alternative Pathways to Teaching, Induction, and Professional Development

Teaching well requires substantial knowledge and skill. The Task Group sought to understand the impact that teacher education has on teachers' effectiveness. The Task Group considered the empirical evidence on four types of professional training:

- *Preservice teacher preparation*: Initial teacher training, conventionally offered in institutions of higher education;
- *Alternative pathways*: Initial teacher preparation offered outside of conventional teacher education programs*;*
- *Induction programs*: Professional support and additional training within the first years of practice; and
- *Professional development*: Ongoing programmatic professional education of practicing teachers.

The Task Group's review of the available research and the rigor of this research highlights the critical need for more and better studies tracing the relationship between specific approaches to teacher education (i.e., curricula, pedagogy and assessment, instructors, structures, and settings) and teachers' capacity for teaching and their students' learning. However, existing research on aspects of teacher education, including standard teacher preparation programs, alternative pathways into teaching, support programs for new teachers (e.g., mentoring), and professional development, is not of sufficient rigor or quality to permit the Panel to draw conclusions about the features of professional training that have effects on teachers' knowledge, their instructional practice, or their students' achievement.

Preservice Teacher Preparation

The Task Group found very few empirical studies that addressed the question of the impact of preparation programs on student achievement or teachers' mathematical content knowledge. Unfortunately, none were of sufficient rigor or quality to allow us to draw conclusions about the relationship of particular features of teacher preparation programs and their effects.

Alternative Pathways Into Teaching

Determining how different types of pathways into teaching may affect the knowledge and skill teachers' bring to their work is a key policy question in a time when issues of teacher recruitment, retention, and quality are paramount. Overall, evidence is mixed on the effects of teachers' pathways into teaching and their relationship to students' achievement. Because of differences in specifications across studies, drawing specific conclusions about alternative pathways in general would be difficult to do from these studies. However, the evidence suggests that the differences among current pathways are small or not significantly different and that variation within programs is much greater than that across programs.

Induction Programs

National reports calling for higher-quality teaching, higher teacher retention rates, and stronger student achievement identify support of new teachers—induction—as an area for improvement. The Task Group found a dearth of peer-reviewed research on early career support programs for mathematics teachers that looked at outcomes related to student achievement or teacher mathematics knowledge; key outcomes for much of the current induction literature is teacher retention, satisfaction, and beliefs. Given the current expansion of induction programs, it is important to assess the effectiveness of induction programs on teachers' effectiveness. Until induction programs are content-specific or include specific content activities, it may be difficult to determine the effectiveness of induction on the mathematical knowledge of teachers or on students' achievement gains.

Professional Development Programs

Practicing teachers continue their teacher education through in-service professional development. Studies that investigated the impact on teacher content knowledge lacked either comparison groups or direct measures of teacher content knowledge, instead relying on teacher self reports. Across studies that investigated the relationship between teacher professional development programs and students' mathematics achievement, there was a positive effect of professional development on student achievement. However, there was insufficient evidence on the specific features of professional development that impact student achievement to make conclusions about what forms of, or approaches to, professional development are effective.

Conclusions and Recommendations Regarding Teachers' Education

Despite common beliefs about qualities of effective teacher education, there is no strong evidence on the relationship between any specific form of teacher education and either teachers' mathematics content knowledge or their students' mathematics learning. Further, there is even less evidence to identify the specific features of training that might account for any program impacts, providing little insight into the crucial components of teacher education.

High-quality research must be undertaken to create a sound basis for the mathematics preparation of elementary and middle school teachers within preservice teacher education, early-career support, and ongoing professional development programs. Outcomes of different approaches should be evaluated, in part, by using reliable and valid measures of mathematical knowledge that are demonstrably associated with gains in student achievement.

Studies are needed with designs that lead to knowledge about the impact of different approaches to professional development, and that permit comparisons with other potential impacts on teacher capacity and effectiveness (e.g., experience, curriculum, curriculum policy). Such research will depend not only on rigorous designs, but also on valid and reliable measures of the key outcome variables: teachers' mathematical knowledge and skill, instructional quality, and student learning.

Key questions on which robust evidence is needed include:

- Does teacher education (including preservice training of different kinds, professional development, and early career induction programs), have an impact on teachers' capacity for teaching and on their students' achievement?

- What are key features of teacher education (e.g., duration, structure, quantity, content, pedagogy, structure, relationship to practice) that have effects on teacher capacity and on student achievement?

- How do contexts (school, students, teachers, policy) affect the outcomes of professional development?

- How do different amounts of teacher education affect outcomes and effects?

Teacher Incentives

Compensation is often cited as a key factor in improving teacher quality, and programs are focusing on compensation as a recruitment and retention strategy. The Task Group investigated the evidence on different salary schemes that aim to recruit, reward, and retain skillful teachers. The programs utilize a variety of labor market incentives including:

- *Skills-based pay:* Paying more to teachers who have technical skills that are in demand in other sectors of the economy, such as teachers with degrees in mathematics;

- *Location pay:* Compensating teachers for working in conditions they view as unfavorable, such as those associated with high-poverty, low-achieving schools; and

- *Performance-based pay:* Paying more to mathematics teachers who are more productive in raising student achievement.

The Task Group examined research on each of these approaches to teacher compensation.

Skills-Based Pay

College students' decisions to prepare for and enter into teaching depend on how the salary structure for teachers compares with those in competing occupations. The magnitude of the salary differential between the private sector and the teaching profession for those who enter teaching with technical training is large, with a moderate difference on entry and a rapidly increasing gap over the first 10 years of employment. Teachers of mathematics and science are significantly more likely to move from or leave their teaching jobs because of job dissatisfaction than are other teachers, and of those who depart because of job dissatisfaction, the most common reason given is low salary.

Location Pay

Research on the effects of location-based pay, intended to attract or retain skilled teachers in schools that serve under-resourced communities, yielded mixed results. The effectiveness of such salary schemes is affected by the size of differential in pay, the gender and experience of the teacher, whether the bonus is a one-time signing bonus or permanent, as well as other factors.

Performance Pay

There is a substantial variability in extant merit-pay systems. The Task Group identified four different dimensions of variability of "merit" pay: whether salary differentials are tied to schools' performance or that of individual teachers, how significant the differential is, the degree to which the scheme is focused on student performance, and whether the plan seems continuous or is a short-term experiment. Across the studies reviewed, each performance pay approach yielded some positive effects on student achievement.

Recommendations and Cautions Regarding Teacher Incentives

The results from research on teacher incentives generally support the effectiveness of incentives, although the methodological quality of the studies in terms of causal conclusions is mixed. The substantial body of economic research in other fields indicating that salary affects the number of workers entering a field and their job performance is relevant. In the context of the totality of the evidence, the Task Group recommends policy initiatives that put in place and carefully evaluate the effects of:

- Raising base salaries for teachers of mathematics to be more competitive with salaries for similarly trained non-teachers;
- Incentives for teachers of mathematics working in locations that are difficult to staff; and
- Opportunities for teachers of mathematics to increase their base salaries substantially by demonstrable effectiveness in raising student achievement.

Knowing more about how various incentive systems affect teachers would enable the design of more effective and efficient incentives.

Elementary Mathematics Specialist Teachers

There have been many calls for the use of "mathematics specialist teachers" at the K–5 level in recent years, with some arguing that the teaching of mathematics even in the elementary grades calls for specialized knowledge, while most elementary teachers are generalists. The Task Group sought to learn what is known about mathematics specialists at the elementary level. To contribute to thoughtful consideration of the issues involved in restructuring teacher roles around the idea of mathematics specialists, the Task Group reviewed a range of models in current use in the United States and abroad, and sought evidence about their effectiveness.

Models of Elementary Mathematics Specialist Teachers

The Task Group identified three models of mathematics specialist teachers: the lead teacher or mathematics coach, who is a resource person for their coworkers and does not directly instruct students; the specialized teacher, who is responsible for the direct instruction of students; and the pull-out mathematics specialist, who directly instructs individuals or small groups of students within a classroom who have been identified as either failing to meet or exceeding their grade-level standards. Mathematics specialists as lead teachers or mathematics coaches are more common than the other two models.

Evidence for Effectiveness of Elementary Mathematics Specialist Teachers

The Task Group identified very few studies that probed the effectiveness of elementary mathematics specialists. Out of 114 potentially relevant pieces of literature, only one explored the effects of specialized mathematics specialists on student achievement in elementary schools. The study found no difference in the mathematics gain scores of students in an elementary school with a departmentalized structure compared to students in a school with a self-contained class structure.

Costs Associated With Mathematics Specialists

The costs associated with the employment, training, or certification of mathematics specialists were considered. One cost has to do with the funding of the personnel involved and depends on the model used: the specialized-teacher model involves only a redistribution of responsibilities among the existing staff, whereas the lead teacher or mathematics coach or the pull-out program requires the hiring of additional teachers. A second cost is that of the additional training needed for teachers to prepare with the specialization to fill these roles. In addition to tuition costs for participating teachers, there are costs associated with developing and operating programs or courses. Simply taking more college-level mathematics courses would not be sufficient in most cases because regular mathematics courses generally do not focus on the mathematics needed for specialization in elementary and middle school.

Mathematics Specialists Internationally

Full-time elementary mathematics teachers are not widely used in most of the countries that produce high levels of student achievement in mathematics. Only three (China, Singapore, and Sweden) deploy such teachers at the elementary level. Elementary teachers in those countries may enter teaching with a stronger background in mathematics, which may be a factor in the success of those countries with mathematics education.

Conclusions Regarding Elementary Mathematics Specialist Teachers

The Panel recommends that research be conducted on the use of full-time mathematics teachers in elementary schools. These would be teachers with strong knowledge of mathematics who would teach mathematics full-time to several classrooms of students, rather than teaching many subjects to one class, as is typical in most elementary classrooms. This recommendation for research is based on the Panel's findings about the importance of teachers' mathematical knowledge. The use of teachers who have specialized in elementary mathematics teaching could be a practical alternative to increasing all elementary teachers' content knowledge (a problem of huge scale) by focusing the need for expertise on fewer teachers.

I. Introduction

Teachers are crucial to students' opportunities to learn and to their learning of mathematics. Substantial differences in the mathematics achievement of students are attributable to differences in teachers. A meta-analysis of the findings from seven large studies of variation in teacher effects found that 11% of the total variability in student achievement gains in mathematics across one year of classroom instruction was attributable to teachers (Nye, Konstantopoulos, & Hedges, 2004). The authors, noting that all the studies in their review were correlational, conducted a new analysis using data from the Tennessee class-size experiment, also known as the Student-Teacher Achievement Ratio (STAR) project. An analysis of STAR data has a methodological advantage over prior studies of natural variation in that both students and teachers were randomly assigned to classes. Nye et al. found that differences in teachers accounted for 12% to 14% of total variability in students' mathematics achievement gains in each of Grades 1, 2, and 3.

In a similar vein, Gordon, Kane, and Staiger (2006), using data from the Los Angeles Unified School District for teachers in Grades 3 through 5, report that the average student assigned to a teacher in his or her 3rd year of teaching who was in the bottom quartile during his or her first 2 years of teaching lost on average 5 percentile points on a mathematics assessment relative to students with similar baseline scores and demographics. In contrast, the average student assigned to a top-quartile teacher gained 5 percentile points relative to students with similar baseline scores and demographics. Thus, the average difference between being assigned a top-quartile or a bottom-quartile teacher was 10 percentile points.

These are large effects, larger, for example, than those that have been shown to result from significant reductions in class size. Important to note is that these effects are for one year of instruction. Teacher effects are much larger when they combine across years of instruction. Sanders and Rivers (1996) used value-added methods to measure the effectiveness of all math teachers in Grades 3, 4, and 5 in two large metropolitan school districts in Tennessee. The growth in academic achievement by students in each teacher's class relative to all other teachers was used to identify the most effective (top 20%) and the least effective (bottom 20%) teachers. The progress of children assigned to these low- and high-performing teachers was tracked over a 3-year period. Children assigned to three effective teachers in a row scored at the 83rd percentile in mathematics at the end of fifth grade, while children assigned to three ineffective teachers in a row scored at the 29th percentile. Using a different methodology, Rivkin, Hanushek, and Kain (2005) came to a similar conclusion: The cumulative effects of children being taught by highly effective teachers can substantially eliminate differences in student achievement that are due to family background.

Unfortunately, little is known about what accounts for these individual differences in teachers' ability to generate gains in student learning. Investigating the best evidence about instructional practices that affect achievement is the province of the Instructional Practices Task Group. The Teachers and Teacher Education Task Group has sought to learn the impact of a teacher's own knowledge on their students' achievement and how teachers can be best recruited, prepared, supported, and rewarded.

The four questions that structured the work of the Teachers and Teacher Education Task Group are:

1) What is the relationship between the depth and quality of teachers' mathematical knowledge, and students' mathematics achievement?

2) What is known about programs that help teachers develop the necessary mathematical knowledge for teaching? Which of these programs have been shown to affect instructional practice and student achievement?

3) What types of recruitment and retention strategies are used to attract and retain highly effective teachers of mathematics? How effective are the strategies?

4) What evidence exists for the effectiveness of elementary mathematics specialist teachers with respect to student achievement? What models exist for elementary mathematics specialists and their preparation?

The Task Group found an uneven research base to address these questions. Included in this report are the Task Group's summary and recommendations, their observations about the nature of the empirical evidence currently available, and recommendations for the research needed to improve practice and policy with respect to teachers. The Task Group's last question, on elementary mathematics specialist teachers, differs from the others as it relies primarily on anecdotal information regarding the effectiveness of mathematics specialists. Despite the paucity of empirical research, the Task Group believes that there is a useful contribution that the Panel can make to debates about this idea. What the Task Group has produced is a synthesis of the programs and models that exist in the United States, some distinctions among different models, and some commentary on the costs and benefits of those different models.

II. Methodology

The Task Group identified and organized the available scientific evidence into several categories as they reviewed studies related to each of the four research questions addressed by the group. The quality of the evidence varied for each research question, as described in the individual sections of the report. For example, because experimental manipulation of teacher characteristics, such as teacher preparation or teacher content knowledge, is not easily done, the Task Group relied on correlational studies investigating the relationship of these variables to student achievement. Following is a discussion of the categories for studies with quantitative designs that were considered for inclusion because they provided the most rigorous evidence available, followed by a discussion of the procedures for identifying relevant research and synthesizing the results. In their report, the Task Group relied primarily on the highest quality studies available for each question.

A. Categories of Studies

1. High-Quality Studies

The high-quality studies identified for the research questions addressed by the Task Group included those with designs that were randomized controlled trials (RCTs), quasi-experimental studies that included baseline equivalence of treatment and control groups formed other than by random assignment, or correlational studies that met stringent quality standards. The strongest quality quasi-experimental and correlational studies were determined on the basis of: 1) the size of sample, 2) appropriate and adequate statistical controls, 3) use of multiple specifications, or tests for robustness of results, or both 4) use of individual-level versus aggregated data, and 5) the appropriateness and strength of the identification of the outcomes variable. These elements are reported for each study reviewed. The correlational studies in this category used such multivariate methods as regression analysis in the form of either standard ordinary least squares (OLS) or hierarchical linear modeling (HLM).

More specifically, the correlational studies in this category all had these three characteristics:

1) A baseline control for the outcomes measure (e.g., prior student mathematics achievement);

2) Use of multivariate regression analysis; and

3) A strong outcome measure (e.g., standardized test to measure student mathematics achievement).

and at least three of the following four characteristics:

1) Sample size—more than 1,000 observations on student mathematics achievement;

2) Statistical controls—contained a pretest control for student mathematics achievement plus controls for other relevant student and teacher characteristics;

3) Multiple specifications—multiple model specifications or other robustness checks on the results; and

4) Student-level data—the analysis was conducted with student-level data that had not been aggregated beyond the classroom of students instructed by a teacher.

Strong quasi-experimental and correlational studies formed the core of the research available to support conclusions and recommendations of the Task Group.

2. Moderate-Quality Studies

Studies in this category were those that were empirical, but did not meet a sufficient number of the standards established for the highest category. For instance, some of the studies in this category are those with strong research designs but weak outcome measures, such as indirect measures of teacher content knowledge or student achievement. All of the studies included in this category used correlational methods, such as regression analysis in the form of either standard OLS or HLM.

3. Lesser-Quality Studies

Studies in this category met very few of the criteria established for the highest quality studies and/or their measure of a variable of interest was extremely weak. For example, studies where the measure for teacher content knowledge was defined as teacher certification in a field other than mathematics or a score on a mathematics test that the teacher took many years earlier fall into this category.

B. Procedures

1. Literature Search and Study Inclusion

Literature searches were conducted to locate studies on the relationship between selected teacher characteristics and student learning in mathematics. Electronic searches were made in PsycINFO and the Social Sciences Citation Index (SSCI) using search terms identified by the Task Group. Studies were also identified through manual searches of relevant journals and reference lists, and recommendations from experts. Abstracts from these searches were screened for relevance to research questions and appropriate study design. For each study that met the screening criteria, the full study was examined to determine whether it met the inclusion criteria specified below. Citations from those articles and research reviews were also examined to identify additional relevant studies.

Criteria for Inclusion:

- Published between 1975 and 2007;
- Involved Grade K–12 students studying mathematics or involved teachers of mathematics for Grade K–12 students;
- Available in English;
- Published in a peer-reviewed journal, government or national report, book, or book chapter; and
- Used multivariate analysis with statistical controls.

2. Effect Size Calculations

Standardized regression coefficients typically were used as effect size measures because they are often the only data available. The equation used to calculate standardized regression coefficients is equation 2.3 from Bring (1994):

$$B_i = \hat{\beta}_i (s_i / s_y)$$

In this equation, $\hat{\beta}_i$ is the standardized regression coefficient on variable *i*, $\hat{\beta}_i$ is the regression coefficient reported in the paper for variable *i*, s_i is the standard deviation of variable *i*, and s_y is the standard deviation of the dependent variable. Thus, to calculate the standardized regression coefficient, the regression coefficient is multiplied by the standard deviation of the independent variable, and divided by the standard deviation of the dependent variable. The first element required for this calculation, $\hat{\beta}_i$, was available in any paper where a regression coefficient on the variable of interest was reported, which included all of the high-quality studies. However, the two required standard deviations were frequently not reported, and in those cases the standardized regression coefficient could not be calculated.

The standardized regression coefficient is interpreted as the change in the dependent variable as a fraction of the standard deviation that results when the independent variable changes by one standard deviation, holding the other variables constant. It is important to note, however, that although this is the best method available to calculate effect sizes given the data available in these papers, it is somewhat controversial and thus should be interpreted with caution. Essentially, the standard deviations used here are not entirely appropriate, because they are the overall standard deviations, which are much larger than the standard deviations when all independent variables but one are held constant. For more detail and an example of why standardized regression coefficients can be misleading, see Bring (1994).

When possible, the Task Group applied the What Works Clearinghouse (WWC) guidelines to calculate standardized mean differences in mathematics achievement.[1] Using C*omprehensive Meta-Analysis, Version 2* software, Hedges *g* standardized mean differences were calculated for these studies. The standardized mean difference is defined as the difference between the mean score for the treatment group minus the mean score for the comparison group, divided by the pooled standard deviation of that outcome for both the treatment and comparison groups.

In cases where schools, teachers, or classrooms were assigned (either randomly or nonrandomly) into intervention and comparison groups, and the unit of assignment was not the same as the unit of analysis, the effect size and accompanying standard error were adjusted for clustering within schools, teachers, or classrooms. This analysis used WWC guidelines to adjust for clustering, applying the default intra class correlation (ICC) adjustment for achievement guidelines of 0.20 when actual ICC values were unavailable.[2]

---

[1] See http://ies.ed.gov/ncee/wwc/twp.asp for the guidelines.
[2] See http://ies.ed.gov/ncee/wwc/overview/review.asp?ag=pi for more information on this issue.

III. Findings

A. Teachers' Knowledge of Mathematics

What evidence exists about the relationship between the depth and quality of teachers' knowledge of mathematics and gains in students' achievement?

1. Introduction

It is widely assumed—some would claim common sense—that teachers must know the mathematical content they teach. Yet verifying this assertion and making it more precise for the purposes of teacher education, professional development, and policy have proved challenging for researchers. How much mathematics course work do teachers need to take? How much do they need to know? What exactly do they need to know, and do they need to learn it? On the one hand, a simple answer—and an obvious one—is that teachers must have a strong grasp of the curriculum that they are responsible to teach. Knowing the curriculum well enough also requires knowledge at levels beyond their grade assignment. On the other hand, showing the effect of different kinds or amounts of teacher mathematical knowledge on student achievement has been far from easy or conclusive. Expert opinion can offer insight into the mathematical demands of teaching.

In this section, the Task Group considered empirical evidence that might help to shape teachers' education and policy: How does the level and scope of teachers' own mathematical attainment affect the learning of their students?

This is, of course, hardly a new question. The relationship between teacher characteristics and student achievement outcomes has been examined extensively over the past four decades, reaching as far back as a study by Coleman in the 1960s (Coleman et al., 1966). This education production function research attempted to determine the relationships between inputs, such as school and teacher characteristics, that can be modified (e.g., class size, expenditures on teacher salaries) and desired education outcomes (e.g., increased student achievement). The findings from this body of work have not produced consistent conclusions (e.g., Hanushek, 1986, 1989; Hedges, Laine, & Greenwald, 1994; Krueger, 1999; Monk, 1992), especially regarding matters of the relationship between education expenditures and student performance. Yet, as Goldhaber (2007) noted, a major finding from this literature was that of all school-related factors, teacher quality dominates effects on student achievement. As such, more recent interest in school inputs has turned to examining how the characteristics of teachers and the policies that affect them (e.g., teacher certification, teacher testing) may be related to student achievement. In this section the Task Group turns to examining how teachers' mathematical knowledge is related to student achievement.

Studying the relationship of teachers' mathematical knowledge to student achievement requires a measure of content knowledge. In the studies the Task Group reviewed and that met their standards of rigor, they identified three ways in which teachers' mathematical knowledge has been measured and related to student achievement gains: teacher certification, mathematics course work, and tests of teachers' mathematical knowledge. Other paths to assess teachers' content knowledge include scores on certification examinations, individual interviews and structured tasks, and teachers' skill with mathematics in the context of actual instruction. The Task Group did not locate studies that examined the effects of measures on teachers' performance on the learning of their students, and so these other approaches are not reflected in their initial summary. Consistent with the findings of other researchers who also have examined this body of literature, the Task Group finds that when studies are combined by type, including those that rely on proxies for teacher knowledge, such as certification status, the general results are mixed (Mandeville & Liu, 1997; Hill, Rowan, & Ball, 2005). When studies that use direct measures of teachers' mathematical knowledge are isolated, the effects are more positive.

The Task Group summarizes the evidence available about teacher knowledge effects, comments on the quality and reliability of that evidence, and suggests a set of warranted conclusions and inferences based on the strongest evidence. In addition, for the strongest group of studies, the Task Group provides the magnitude of the findings for illustrative purposes.

2. Teacher Certification as a Measure of Mathematical Content Knowledge

The Task Group begins this synthesis by examining literature that uses teacher certification as a proxy for teachers' mathematical content knowledge. A drawback to certification status is its inexactness as a measure of teachers' content knowledge. This stems from several problems. The first is one of self-selection. If a teacher's certification status is correlated with other characteristics that may also affect student outcomes (e.g., motivation), then any impact of certification on student achievement may not necessarily reflect the effect of greater mathematical content knowledge. The second problem is that, to the extent that certification status does not measure more nuanced elements of teacher mathematical content knowledge, if and where those "undetectable" teacher-to-teacher differences exist and are significant determinants of student outcomes, estimates of the effect of teacher mathematical content knowledge will be less precise, i.e., biased toward finding no impact on student outcomes. Another limitation is that different types of certification status (e.g., standard versus emergency) complicate the understanding of the effect of teacher mathematical content knowledge, primarily because how "true" mathematical content knowledge differs between these two groups of certified teachers is not completely clear (see Darling-Hammond, Berry, & Thoreson, 2001; Fetler, 1999). Finally, the small percentage of uncertified teachers may also raise questions about what is being measured.

a. Overall Findings

Overall, findings about the impact of teacher certification on student achievement in mathematics have been mixed, even among the most rigorous and highest-quality studies. Research in this area has not provided consistent or convincing evidence that students of teachers who are certified in mathematics gain more than those whose teachers are not. This may be in part due to some of the drawbacks mentioned above (Goldhaber & Brewer, 2000; Rowan, Correnti, & Miller, 2002; Fetler, 1999; Mandeville & Liu, 1997).

As can be seen in Table 1, three studies found positive effects of teacher certification on student achievement (Goldhaber & Brewer,[3] 1997b, 2000; King Rice, 2003*)*. Interestingly, each made use of the U.S. Department of Education's National Educational Longitudinal Survey of 1988 (NELS:88), using as the dependent measure gain scores in student achievement, or student achievement with prior grade achievement as a control. The earlier work by Goldhaber and Brewer (1997b) examined differences in student achievement in mathematics in 10th grade with scores in 8th grade as a control, while their 2000 study and King Rice's 2003 research examined 12th-grade differences in student achievement with scores in 10th grade as a control. Each of these studies used scores on the NELS:88 mathematics test as its dependent measure, and each used a number of statistical controls in their regression analyses, including prior mathematics test scores. And though their point estimates of the effect of certification differ, all found positive impacts of teacher certification in mathematics on student achievement.

Conversely, two of the high-quality studies and one of moderate quality found no significant effect of teacher certification as a predictor of student achievement in mathematics. These include studies by Kane, Rockoff, and Staiger (2006); Hill et al. (2005); and Rowan et al. (2002).[4] Each of these studies made use of distinctly different data sets that vary in terms of measures (for both student test scores and sample characteristics) and sample representation. Specifically, Kane et al. used statewide data and state-specific standardized test scores for mathematics in New York City, while the other two studies used data spanning states that were originally collected as part of another study (Hill et al.; Rowan et al.).[5] Across all of these data sources, none of the studies found a significant effect of teacher certification status on students' mathematics achievement. Thus, the strongest evidence currently available about the effect of teacher certification on student achievement in mathematics (i.e., the high- and moderate-quality studies reviewed for this section) is mixed.

[3] A third study (Goldhaber & Brewer, 1997a) uses the same data set and sample of students and teachers as the 1997b study. As the latter study includes additional predictors (i.e., teacher classroom behavior) the Task Group chose to include it as it appears to be a better specified model.

[4] Although all three of these studies represent solid empirical investigation, the Task Group considers the models in the Rowan, Correnti, and Miller (2002) study to be of *moderate* quality due to data limitations that the authors faced in their analysis. Specifically, the authors acknowledged that their results and subsequent claims about the effect of certification reflected estimates based on less than 6% of their sample. The models reported in the Harris and Sass (2007a) study investigated the effect of National Board for Professional Teaching Standards (NBPTS) certification, which is very different than standard certification. Thus, while the Task Group still considers these reports in drawing their conclusion, they represent less powerful evidence to support the argument at hand.

[5] Hill, Rowan and Ball (2005) make use of a study of instructional improvement initiatives in schools. Rowan, Correnti, and Miller (2002) uses the *Prospects: The Congressionally Managed Study of Educational Growth and Opportunity 1991–1994.*

Table 1: Quality Characteristics of Models Examining Impact of Teacher Certification on Student Achievement in Mathematics, by Study and by Overall Quality

Authors	Sample Size	Pretest Control[a]	Other Controls				Multiple Specifications	Student-Level Analysis	Identification Measure	Grade(s)
			Student	Teacher	Class, School, or District	Family				
High-Quality Studies										
Goldhaber, & Brewer, 1997b	5,149 students, 2,245 teachers 638 schools 3498 math classes	X	X	X	X	X	X	X	Certified in math	10
Goldhaber, & Brewer, 2000	3,786 students, 2,098 teachers	X	X	X	X	X	X	X	Level of certification[b]	12
King Rice, 2003	3,696 teachers	X		X	X		X		Certified in math	12
Kane, Rockoff, & Staiger 2006	1,462,100 student-year observations	X	X	X	X	X	X	X	Type of certification	Panel data: 3–8
Hill, Rowan, Ball, 2005	2,963 students, 699 teachers	X	X	X	X	X	X	X	Certified	1 and 3
Moderate-Quality Studies										
Rowan, Correnti, & Miller, 2002	Panels of about 4,000 students in more than 300 classrooms and more than 120 schools.[e]	X	X	X	X	X		X	Certified in math[c]	Cohort 1: 1–3; Cohort 2: 3–6
Harris, & Sass, 2007a	1,112,984 student-year observations	X	X	X	X	X	X	X	NBPTS certification[d]	4–10
Lesser-Quality Studies										
Darling-Hammond, 1999	44 NAEP state averages for students, 52,000 public school teachers, 9,500 public schools, 5,6000 school districts								Well-qualified: state certification + equivalent of math major (B.A. or M.A.)	4 and 8
Fetler, 1999	921,437 Total Grade 9: 347,201 Grade 10: 313,303 Grade 11: 260,933			X	X	X			Emergency certification, school-level	9–11
Hawkins, Stancavage, & Dossey, 1998	Unweighted - Grade 4: 6,627; Grade 8: 7,146; Grade 12: 6,904; Weighted - Grade 4: 3,714,998; Grade 8: 3,570,116; Grade 12: 2,830,443								Certified in math, certified in education	4, 8, and 12
Larson, 2000	6,474 students, 185 teachers	X							Certified in math	1st semester algebra students
Mandeville, & Liu, 1997	203 teachers				X				Grade level certification	7

[a] Includes use of pretest as control variable, gain scores, or value-added models.
[b] Teachers responded to a question about their certification in mathematics or science. Based on their responses, they were classified as one of five groups: standard certification (reference group), probationary certification, emergency certification, private school certification, and no certification.
[c] The authors caution that only 6% of the sample had special certification, subject-matter degrees, or both.
[d] National Board for Professional Teaching Standards (NBPTS) certification can be general or subject-specific—this measure includes both types.
[e] The authors conducted 12 analyses, given the design of the parent study *(Prospects: The Congressionally Managed Study of Educational Growth and Opportunity 1991–1994)* the number of students, classrooms, and schools varies over time as the pathways students take diverge.

b. Strength of the Findings

The effect of teacher certification status on students' mathematics achievement remains somewhat ambiguous. Of the 12 studies the Task Group reviewed on this subject, 5 provide the highest-quality evidence due to five features of their design: 1) sample size, 2) appropriate and adequate statistical controls, 3) multiple specifications or tests for robustness of results,[6] 4) micro-level versus aggregated data, and 5) the appropriateness and strength of the identification of teacher content knowledge.[7] These elements are reported in Table 1 for the 12 studies that use teacher certification status as a proxy for mathematical content knowledge.

Seven studies (those in the first two panels of Table 1, the high- and moderate-quality studies) used regression analysis in the form of either standard OLS or HLM. Although the other five studies (in the bottom panel of the table) support the general conclusions of the Task Group's findings, they represent the weakest evidence because they lack important qualities such as adequate controls (e.g., pretest scores for students), highly detailed data sets, or meaningful alternative specifications. The implications of these studies, therefore, must be interpreted with caution.

c. Magnitude of the Findings

Although most of these studies use very different measures of student achievement to assess the effect of teacher certification, it may be beneficial to know something about the magnitude of the effects reported in each piece. Already known are the different measures of dependent variables that can complicate the ability to compare regression coefficients across studies. Additionally, many studies lack sufficient information to transform measures of impacts (e.g., regression coefficients, mean differences) into standardized measures across studies. As a result, this Task Group can do little more than report standardized regression coefficients where data are available, and be cautious in their interpretation and comparison of these effects, regardless of their statistical significance. Table 2 lists the reported impacts of teacher certification on student achievement for the high-quality studies in terms of standardized regression coefficients.

---

[6] Robust results are those that are consistent even when multiple specifications or models are used.

[7] Of the five studies that met the criteria for high quality, only one (King Rice, 2003) does not employ student- and family-level controls in its analysis. This is because King Rice aggregates her student- and family-level controls to the classroom (average) level. Also, the use of a pretest control in conjunction with other classroom- teacher- and school-level controls suggests her analysis satisfies the criteria of appropriate and adequate statistical controls.

Table 2: Reported Impacts of Models Examining the Effect of Teacher Certification on Student Achievement in Mathematics for the High-Quality Studies

Authors	Dependent Measure	Standardized Regression Coefficient[b]	Analytic Technique
High-Quality Studies[a]			
Goldhaber, & Brewer, 1997b	NELS test battery scores (Grade 10)	0.06*	Generalized Least Squares
Goldhaber, & Brewer, 2000	NELS test battery scores (Grade 12)		Ordinary Least Squares
	Probationary certification in subject[c]	0.01	
	Emergency certification in subject[c]	0.00	Ordinary Least Squares
	Private school certification[c]	-0.01	Ordinary Least Squares
	No certification in subject[c]	-0.01*	Ordinary Least Squares
King Rice, 2003	NELS test battery scores (Grade 12)	0.02*	Ordinary Least Squares
Kane, Rockoff, & Staiger, 2006	NYC standardized test scores (value-added Grades 3–8)	0.00	Ordinary Least Squares
Hill, Rowan, & Ball, 2005	Terra Nova scores		Hierarchical Linear Modeling
	- Grade 1	0.00	
	- Grade 3	0.00	

*p < .05
[a] High-quality studies are defined on page 3.
[b] The standardized coefficient was calculated for those studies with sufficient data using the formula provided in Bring (1994) $B_i = \hat{\beta}_i(s_i / s_y)$
[c] The comparison is to standard certification in subject.

As shown in Table 2, where there is a significant effect of teacher certification in mathematics on student achievement, it is quite small. Also, because teacher certification in mathematics is a remote proxy for mathematical content knowledge, this measure does not allow strong inferences about the effect of teachers' knowledge on their students' achievement.

3. Content Course Work and Degrees as Measures of Mathematical Content Knowledge

Mathematics course work and field-specific degrees are a second common proxy for teachers' mathematical content knowledge. Although these are different, they both focus on teachers' completion of college-level mathematics study. Consequently, these measures frequently appear together within the same data set and, thus, are often jointly considered within the same analysis. Therefore, for this section the Task Group discusses and synthesizes the literature and subsequent evidence presented for the relationship between these measures of teacher mathematical content knowledge and student achievement.

Although the amount of course work or the possession of a degree in mathematics are both closer predictors of a teacher's mathematical knowledge than certification status, these are still both proxies for that knowledge and each has unique validity problems. Neither measures the actual command of specific mathematical topics and skills. Neither measures what an individual actually learned, which may vary substantially from person to person. There is similarly no information about the correspondence between particular courses and the school curriculum for which teachers are responsible. Thus, as a measure of the knowledge on which teaching depends, course work or degree attainment may or may not correspond to what teachers use in the course of their work.

Also, as with certification status, there is no guarantee that using the amount of course work completed or type or level of degree circumvents the problems of selection bias previously mentioned. It could be that teachers who engage in generous amounts of mathematics course work or obtain mathematics degrees are particularly motivated to teach mathematics, or possess some other unobservable characteristics unrelated to course work and degree that make them especially effective at teaching mathematics. Finally, these measures do not take into account the passage of time. On the one hand, this may be important if individuals forget lessons from their schooling and change the core of their mathematics instruction over time to satisfy student or school needs or demands (e.g., switching from teaching geometry to teaching calculus).[8] On the other hand, if teachers constantly practice the content-specific skills they need, and those needs do not significantly change over time, then the time lag is decidedly less important.

a. Overall Findings

Much like the research using certification status as a proxy for teachers' mathematical knowledge, the findings in the literature on the impact of content-specific course work and degrees are mixed.[9] Among the seven studies of high quality that examine the impact of a teacher's mathematics course work, or degree, or both, on students' mathematics achievement, several of which look at multiple measures of course work, or degree, or both, four provide estimates of the relationship between degree and student achievement (three are positive; one negative), four provide estimates of the relationship between course work and student achievement (all positive), and two examine the relationship between college credits and student achievement (one positive, one negative). The evidence from this set of studies is somewhat more consistent than the evidence for an effect of teacher certification, thus pointing to a likely positive relationship between teacher content course work and degrees, and student mathematics achievement. Left unexamined are important details, such as how many courses, or which degrees or programs of study make the most difference. Moreover, of the studies that found a positive impact of teacher course work and degrees on student mathematics achievement, many use NELS:88 data and most focus on high school students; thus their scope is limited.

In particular, two reports by Goldhaber and Brewer (1997b, 2000) found positive impacts of course work attainment as noted previously. Both of these studies use the NELS:88 to measure student mathematics achievement, and their samples include students in 10th (1997b) or 12th grades (2000). Findings from both studies indicate that students who have teachers with degrees in mathematics perform significantly higher on the NELS test battery than do students of teachers who are not qualified in math. Rowan, Chiang, and Miller (1997) also use the NELS data for 10th-grade students, finding that teacher's possession of a mathematics degree is associated with higher student achievement. Monk and King (1994), on the other hand, use Michigan State University's Longitudinal Study of American Youth (LSAY), which uses items from the U.S. Department of Education's National Assessment of Educational Progress (NAEP) achievement test in mathematics. That

[8] Though including teacher experience as a control variable (common in the literature) helps to alleviate the inexactness of course work and degrees as measures over time, estimates of the interaction between teacher experience and course work and degrees (not common in the literature) would help to determine actual effects of course work and degrees.
[9] Supported by Wayne and Youngs (2003).

study also focuses on secondary school students but examines teacher course work at the graduate and undergraduate level. They found a positive relationship between number of mathematics courses taken and student achievement. The positive effects of teacher preparation in mathematics thus seem to exist across different data sources, although as noted, these studies all focus on high school grades.

Less clear is how teachers' level of conventional college mathematics study affects student achievement below ninth grade. Rowan et al. (2002) used a survey that followed two cohorts of students, starting in Grades 1 and 3, for 4 years. The authors found a negative impact for an advanced mathematics degree on student achievement, but acknowledge that fewer than 6% of teachers (approximately 43 total) had subject-matter specific degrees. Harris and Sass (2007b) examined a number of different measures for students of all different ages, and although they mostly found no effects, they identified positive impacts for some measures and negative for others. Hill et al. (2005) found that teachers' mathematics course work did not significantly predict gains in student achievement.

Thus, the results across the studies that use teachers' attainment as a proxy for mathematical knowledge are mixed. The strongest suggestion is that the level of teachers' mathematics study may predict student achievement at the high school level. Evidence was not uncovered to support this relationship below ninth grade. It may be that using course work as a proxy for teachers' actual knowledge is a less valid measure at this level than it is at the secondary school level where the content teachers teach is closer to the content they study in college.

b. Strength of the Findings

As described earlier in the section on certification, the Task Group chose to focus on a limited number of studies of the highest quality, based on 1) sample size, 2) appropriate and adequate statistical controls, 3) multiple specifications or tests for robustness of results, 4) micro-level versus aggregated data, and 5) the appropriateness and strength of the identification. These elements are reported in Table 3. Each of the studies used detailed regression analysis, in the form of either standard OLS or HLM.

Studies of lesser quality are also reported in Table 3. While these studies support the general conclusions of the Task Group's findings, they represent weaker evidence because they lack such important qualities as adequate controls (e.g., pretest scores for students), highly detailed data sets, and meaningful alternative specifications. Therefore, the implications of this research must be interpreted with caution.

c. Magnitude of the Findings

Although these studies use differing measures of student achievement and of teacher mathematics course work and degrees, there is value in knowing something about the magnitude of the effects reported in each piece. Table 4 lists the reported impacts of teacher course work and degrees on student achievement for the highest quality studies using standardized regression coefficients to allow comparisons across the studies where data are available. Table 4 shows that for teacher certification, the impact of teacher course work and degrees on student mathematics achievement, when a standardized coefficient can be calculated, is quite small. This finding is consistent across measures and grade levels, although few studies have sufficient data for calculating the magnitude of the effect. Therefore, generalizations are not appropriate here.

Table 3: Quality Characteristics of Models Looking at Impact of Teacher Mathematics Course Work and Degrees on Student Achievement in Mathematics, by Study and by Overall Study Quality

			Other Controls							
Authors	**Sample Size**	**Pretest Control**[a]	**Student**	**Teacher**	**Class, School, or District**	**Family**	**Multiple Specifications**	**Student-Level Analysis**	**Identification Measure**	**Grade(s)**
High-Quality Studies										
Goldhaber, & Brewer, 1997b	5,149 students, 2,245 teachers, 3,498 math classes, 638 schools	X	X	X	X	X	X	X	B.A. and graduate degree in math	10
Goldhaber, & Brewer, 2000	3,786 students, 2,098 math teachers	X	X	X	X	X	X	X	B.A. and graduate degree in math	12
Harris, & Sass, 2007b	Grades 4–5: 514,620 students/ 785,780 observations; Grades 6–8: 542,289 students/ 784,423 observations; Grades 9–10: 426,474 students/ 667,698 observations Teachers: Grades 3–5: 2,134; Grades 6–8: 943; Grades 9–10: 639	X	X	X	X	X	X	X	Course work in math, college credits, math degree	4–10
Hill, Rowan, & Ball, 2005	2,963 students, 699 teachers	X	X	X	X	X	X	X	Course work in math content and math methods	1 & 3
Monk, 1994	2,829 students, 608 math teachers	X		X	X		X	X	College degree in math, college courses in math	10 & 11
Monk, & King, 1994	2,831 students, their teachers	X	X	X	X	X	X	X	College courses in math	Cohort 10–12
Rowan, Chiang, & Miller, 1997	5,381 students	X	X	X	X	X	X	X	College degree in math	10
Moderate-Quality Studies										
Rowan, Correnti, & Miller, 2002	Panels of about 4,000 students in more than 300 classrooms and more than 120 schools	X	X	X	X	X	X	X	College degree in math[b]	Cohort 1: Grades 1–3; Cohort 2: Grades 3–6
Lesser-Quality Studies										
Darling-Hammond, 1999	44 state averages for students, 52,000 public schools, 65,000 teachers								Well-qual: certification + math major	4 & 8
Darling-Hammond, Berry, & Thoreson, 2001	3,786 students, 2,078 teachers	X		X				X	College degree in math (BA or MA)	12
Eisenberg, 1977	807 students, 28 teachers		X					X	College math GPA, courses taken	Junior high Algebra I students

[a] Includes use of pretest as control variable, gain scores, or value-added models. Gain scores are the calculated gains of students across a school year.
[b] The authors also caution that they identified the effect of mathematics college degree from a maximum of 6% of the teachers in their sample.

Table 4: Reported Impacts of Models Examining the Effect of Teacher Math Course Work and Degrees on Student Achievement in Mathematics, by Study and By Overall Quality

		Standardized[a] Regression Coefficients			
Authors	**Dependent Measure[b]**	**College Credits**	**Degree/Major**	**Course Work**	**Analytic Technique**
Goldhaber, & Brewer, 1997b	NELS test battery scores (Grade 10)		0.03*		Ordinary Least Squares
Goldhaber, & Brewer, 2000	NELS test battery scores (Grade 12)		Not available		Ordinary Least Squares
Harris, & Sass, 2007b	FCAT scores				2-stage Ordinary Least Squares
	- Elementary		Not available		
	- Middle		Not available		
	- High		Not available		
Hill, Rowan, & Ball, 2005	Terra Nova scores				Hierarchical Linear Modeling
	- Grade 1	0.02			
	- Grade 3	0.05			
Monk, 1994	Scores on NAEP-based tests				Ordinary Least Squares
	- Grade 10		Not available	Not available	
	- Grade 11		Not available	Not available	
Monk, & King, 1994	Gain scores on NAEP-based tests (Grades 10, 11, and 12)			.09*	Ordinary Least Squares
Rowan, Chiang, & Miller 1997	NELS test battery scores (Grade 10)		Not available		Hierarchical Linear Modeling

*$p < .05$
[a] The standardized coefficient was calculated for those studies with sufficient data using the formula provided in Bring (1994) $B_i = \hat{\beta}_i(s_i / s_y)$.
[b] Of the studies that consider measures of mathematics course work as their independent variable, only one (Monk, 1994) partitions its measure by the type of course work, in this case graduate and undergraduate course work. Along these measures, the authors find that there is no significant effect (at the 5 % level) of graduate mathematics course work on student achievement, and that there is a significant effect of undergraduate mathematics course work, but only for one (i.e., the juniors) of the two grades tested.

4. Test Scores and Ad Hoc Measures as Measures of Mathematical Content Knowledge

Using test scores and more proximal measures of teacher mathematical content knowledge allows closer examination of the effect that mathematical knowledge has on student achievement. Such measures escape some of the traditional problems of selection bias and inexactness present in other more conservative measures. These types of measures make it possible to probe more directly the causal link between mathematical knowledge and student achievement. One caution with this approach is that specially developed tests or measures, although intuitively appealing or apparently more relevant, have often not been validated or otherwise checked for their psychometric quality.

a. Overall Findings

Research that has used teacher test scores and other ad hoc measures has also produced mixed results. The Task Group's inability to draw solid conclusions from this literature is in part due to the general lack of quality measures of mathematics content knowledge, as well as the absence of an adequate number of high-quality studies using these types of measures. Overall, the Task Group identified five studies that both examined the effect of mathematical content knowledge (as measured by test scores and other instruments) on student achievement and that met their standards for high or moderate quality. Of these, two studies found a positive and significant effect of mathematical content knowledge on student achievement, another found a positive but statistically insignificant effect, and two other studies found ambiguous effects for various measures with their sample. Unlike the studies reviewed earlier on teacher certification and course work that mostly examined student achievement at the high school level, the studies in this group are focused at the elementary level, making comparisons with other findings difficult.

Clotfelter, Ladd, and Vigdor (2007) examined the relationship of teacher test scores to student mathematics achievement in North Carolina. Although the test did not focus solely on mathematics content, they found that higher teacher test scores are a significant predictor of higher student achievement. Hill et al. (2005) used test items specifically designed to measure the mathematical content knowledge used in teaching and controlled for content knowledge in teaching reading. They found that the measure of content knowledge of mathematics is a significant and positive predictor of student success in math.

Harbison and Hanushek (1992) examined the effect of teacher mathematics test scores on fourth-grade tests on student achievement in Brazil; the authors found a positive effect of teacher test score on student achievement: "At fourth grade, a ten-point improvement in the mean teacher's command of her mathematics subject matter … would engender a five-point increase in student achievement; this is equivalent to a 10% improvement over the mean scores of fourth graders" (p. 114). The effects are not significant at the traditional .05 statistical level.

Two other studies used standardized tests as measures of teachers' knowledge of mathematics. These show ambiguous results for determining the impact of teacher content knowledge on achievement. One study (Harris & Sass, 2007b) used teachers' quantitative Scholastic Aptitude Test (SAT) scores taken at the time of college entry, a measure that is substantially more distal than those used in the studies reported above, to assess the impact of teacher content knowledge on math achievement for students taking the Florida Comprehensive Assessment Test (FCAT). Further, they partitioned their sample by grade to differentiate potential impacts at various grade levels. Ultimately they found that there is no significant effect of teachers' previous higher SAT scores on elementary and middle school students, and that there is actually a negative impact on achievement for high school students.

Mullens, Murnane, and Willett (1996) used teacher test scores on the Belize National Selection Exam (BNSE) to analyze their effect on students' understandings of basic and advanced concepts in math. The authors found that while teacher test scores are not significant predictors of achievement for basic concepts in mathematics, they do exert a positive influence on student understanding of advanced concepts.

b. Strength of the Findings

The Task Group's conclusions about the impact of teacher content knowledge on student achievement, as measured by tests or specially designed measures, could only be drawn from a small number (*n* = 3) of the high-quality studies. Still, evidence from the lesser-quality studies is consistent, pointing in a positive direction. For example, Rowan et al. (1997) identified teacher content knowledge using a teacher's response to a single school-relevant mathematics question and found positive effects for that one item. Additionally, two studies (Harris & Sass, 2007b; Mullens et al., 1996) also are considered moderate quality due to their choice of measurement of teacher mathematical content knowledge. It could be argued that the use of quantitative SAT scores (Harris & Sass) may not capture important elements of mathematical content knowledge that are acquired at the collegiate level. By the same token, the use of eighth-grade BNSE scores (Mullens et al.) is even more likely to be a less relevant measure. Nevertheless, this Task Group cannot discount the possibility that these measures may assess mathematical content knowledge relevant for elementary (or even middle) school teaching. Overall, the evidence here (Table 5), based on a small number of studies, does point more strongly in the direction of a relationship between teachers' usable knowledge of mathematics and students' achievement than do the other measures of teacher content knowledge.

c. Magnitude of the Findings

Similar to the other sections, the Task Group reports the magnitude of the findings for each study in this strand using standardized regression coefficients, where data are available (see Table 6 later in this section). Since each one of these studies uses a distinctly different dependent measure, and only two studies have sufficient data for making comparisons, the Task Group cannot make any meaningful relative interpretations of these findings. However, it should be noted that even though the magnitude of the findings where they are available is quite small, the Hill et al. (2005) findings are substantially larger than the relationships noted for other measures of teacher content knowledge.

Hill et al. (2005) also examined the effect of teacher certification in mathematics and mathematics education course work, in addition to their specific content knowledge measure These comparative analyses show that the specific measure of mathematical knowledge for teaching used by Hill et al. (2005) is measuring a type of understanding and skill that is not captured by certification status or course work measures.

Table 5: Quality Characteristics of Models Looking at Impact of Teacher Test Scores And Other Ad Hoc Measures on Student Achievement in Mathematics, by Study and By Overall Study Quality

			Other Controls							
Authors	**Sample Size**	**Pretest Control***	**Student**	**Teacher**	**Class, School, or District**	**Family**	**Multiple Specifications**	**Student-Level Analysis**	**Identification Measure**	**Grade(s)**
High-Quality Studies										
Clotfelter, Ladd, & Vigdor, 2007	Nearly 1 million student-year observations	X	X	X	X	X	X	X	North Carolina teacher test scores	3–5
Harbison, & Hanushek, 1992	Over 2,500 students	X	X	X	X	X			Brazilian math test scores	2–4
Hill, Rowan, & Ball, 2005	2,963 students, 699 teachers	X	X	X	X	X	X	X	CKT-M test measure	1–3
Moderate-Quality Studies										
Harris, & Sass, 2007b	Grades 4–5: 514,620 students/785,780 observations; Grades 6–8: 542,289 students/784,423 observations; Grades 9–10: 426,474 students/ 667,698 observations Teachers: Grades 3–5: 2,134; Grades 6–8: 943; Grades 9–10: 639	X	X	X	X	X	X	X	SAT Quant. Score	4–10
Mullens, Murnane, & Willett, 1996	1,043 students	X	X	X	X	X		X	BNSE Math Score	3
Lesser-Quality Studies										
Rowan, Chiang, Miller, & 1997	5,381 students	X	X	X	X	X	X	X	A single-item teaching-relevant and difficult math test	10
Sheehan, & Marcus, 1978	1,836 students, 119 teachers	X		X	X				WCET scores	1

* Includes use of pretest as control variable, gain scores, or value-added models. Gain scores are the calculated gains of students across a school year.

Table 6: Reported Impacts of Models Examining the Effect of Teacher Test Scores and Other Ad Hoc Measures on Student Achievement in Mathematics, by Study

Authors	Dependent Measure	Standardized Regression Coefficient[a]	Analytic Technique
High-Quality Studies			
Clotfelter, Ladd, & Vigdor, 2007	North Carolina standardized test (Grades 3–5)		Ordinary Least Squares
	- Score Level	Not available	
	- Score Gain	Not available	
Harbison, & Hanushek, 1992	Brazilian math test (from EDURURAL[b])		Ordinary Least Squares
	- Grade 2	Not available	
	- Grade 4	.02	
Hill, Rowan, & Ball, 2005	Terra Nova scores		Hierarchical Linear Modeling
	- Grade 1	.06*	
	- Grade 3	.05*	

*p < .05
[a] The standardized coefficient was calculated for those studies with sufficient data using the formula provided in Bring (1994) $B_i = \hat{\beta}_i(s_i / s_y)$.
[b] EDURURAL is the research project that emerged from an effort to improve educational performance in rural northeast Brazil.

5. The Mathematical Content and Nature of Teacher Licensure Exams

Recent research treating teacher licensure as a proxy for teachers' mathematical content knowledge has not consistently or convincingly shown that students of teachers who are licensed in mathematics gain more academically than those whose teachers are not (Goldhaber & Brewer, 1997a, 1997b, 2000; Hill et al., 2005; Kane et al., 2006; King Rice, 2003; Rowan et al., 2002). However, teacher licensure exams still play an important role in determining the quality and quantity of math teachers available for employment in schools. To assess the quality of teacher licensure exams, it is first necessary to ascertain the mathematical integrity of the exam questions as well as the relevance of these questions to teaching in the classroom of an elementary or middle school. Such precise information turns out to be difficult to obtain.

Most states use a teacher licensure exam that yields a score as a measure of a candidate's achievement in the subject of mathematics. Each state has designed its own unique system of licensure by choosing among different exams and determining cut scores for them.[10] The *Praxis Series* of exams created by Educational Testing Service (ETS) is the most commonly used teacher licensure exam. Overall, one or more of these exams are currently required in 38 states[11] and the District of Columbia, with 7 of these states[12] using

[10] Connecticut, Kansas, and Missouri use an identical set of certification exams and cut scores for the subject of math, as do Oklahoma and Nebraska.
[11] These are Alabama, Alaska, Arkansas, California, Connecticut, Delaware, Georgia, Hawaii, Idaho, Indiana, Kansas, Kentucky, Louisiana, Maine, Maryland, Minnesota, Mississippi, Missouri, Nebraska, Nevada, New Hampshire, New Jersey, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Utah, Vermont, Virginia, Washington, West Virginia, and Wisconsin.

the *Praxis Series* in conjunction with their own exam. Of the 12 states that do not use *Praxis*, 9[13] use their own licensing exam exclusively, while the other 3 either do not require an exam for licensure or do not use an exam that yields an independent score in the subject of math.[14]

The *Praxis Series* is composed of two separate exams, The *Praxis I* and *II*. The *Praxis I* exams, or Pre-Professional Skills Tests (PPST), are designed to measure basic skills in reading, writing, and mathematics. Most ETS states currently require the *Praxis I* tests for licensure, and often for admission into their teacher education programs. The cut scores currently required for licensure across these states range from 169 to 178 while the average performance range of teacher candidates is from 175 to 183.[15] The *Praxis II* exams in mathematics measure specifically mathematical knowledge and teaching and are not nearly sufficient to warrant the formulation of any conclusions about the licensure of middle school teachers in mathematics.

ETS shared data regarding two *Praxis II* exams in Series 0061 and 0063, one testing mathematical content knowledge, and the other addressing mathematical proofs, models, and problems (both with item performance data). At least one of these two series is used by 33 states as part of the general licensure requirements in mathematics,[16] but 30 of these states[17] also require additional exams, and 24 of these[18] also use a *Praxis II* exam designed specifically for testing mathematical content knowledge at a middle school level.[19] The cut scores in each state for these particular exams supplied by ETS are not known because these exams have been retired by ETS. There is also another unresolved issue. As these exams are for single-subject certification, it would be necessary to know the precise role the exams played in certifying middle school teachers in the 30 states mentioned earlier before any statement can be made about the appropriateness of these exams for the certification of middle school teachers. For example, in the multiple choice exam on content knowledge, 8 of 25 exam items are on high school mathematics (e.g., calculus, trigonometry, conditional

[12] These are Alabama, Alaska, California, Georgia, Indiana, South Carolina, and Washington.
[13] These are Arizona, Colorado, Florida, Illinois, Massachusetts, Michigan, New Mexico, New York, and Texas.
[14] Wyoming requires teacher candidates to demonstrate knowledge of the U.S. Constitution and the Wyoming Constitution either through course work or an exam. Montana does not require any testing for licensure, and Iowa does not use an exam with a separate math score.
[15] The average performance range encompasses the scores earned by the middle 50% of the examinees taking the test (n = 93,805). This statistic provides an indication of the difficulty of the test. The standard error of measurement for this figure (2.4) is a statistic that is often used to describe the reliability of the scores of a group of examinees.
[16] These are Alabama, Alaska, Arkansas, Connecticut, the District of Columbia, Georgia, Hawaii, Idaho, Indiana, Kansas, Kentucky, Louisiana, Maine, Maryland, Minnesota, Mississippi, Missouri, Nevada, New Hampshire, New Jersey, North Dakota, Ohio, Oregon, Pennsylvania, South Carolina, South Dakota, Tennessee, Utah, Vermont, Virginia, Washington, West Virginia, and Wisconsin.
[17] These are Alabama, Alaska, Arkansas, Connecticut, the District of Columbia, Georgia, Hawaii, Idaho, Indiana, Kansas, Kentucky, Louisiana, Maine, Maryland, Minnesota, Missouri, Nevada, New Hampshire, New Jersey, North Dakota, Ohio, Oregon, Pennsylvania, South Carolina, South Dakota, Tennessee, Vermont, Virginia, Washington, and West Virginia.
[18] These are Alabama, Alaska, Connecticut, the District of Columbia, Idaho, Kansas, Kentucky, Louisiana, Maine, Maryland, Minnesota, Missouri, New Hampshire, New Jersey, North Dakota, Ohio, Oregon, Pennsylvania, South Dakota, Tennessee, Vermont, Virginia, Washington, and West Virginia
[19] Here the District of Columbia has been counted as a state and currently uses the *Praxis II* exam designed specifically for testing mathematical content knowledge at a middle school level in addition to an exam testing mathematical content knowledge and one addressing mathematical proofs, models, and problems.

probability, matrices) while only 5 are related to fractions, the central topic of middle school mathematics; of these 5, only 2 (one on percent, the other on rate) directly probe teachers' understanding of fractions. This exam, then, does not seem to assess directly the content for which middle school teachers are responsible.

To determine whether the tests adequately measure teacher content knowledge, precise information about their licensure is needed. At the moment, the collection of such information is greatly hampered by confidentiality issues. The Task Group recommends that there be more openness in this area to facilitate the kind of academic inquiries that are a prerequisite to progress.

a. Implications From the Empirical Evidence

Overall, across the studies reviewed, the signal is that teachers' knowledge of mathematics is a positive factor in students' achievement. However, despite the common-sense nature of this claim, solid evidence about the relationship of teachers' mathematical content knowledge to students' mathematics achievement remains uneven and has been surprisingly difficult to produce. One main reason has been the lack of valid and reliable measures of teachers' mathematical knowledge. The literature has been dominated by the use of proxies for such knowledge, such as certification status and course work. A second reason for the inconsistent findings has been weak study designs. Too few studies set up proper comparisons or use sufficient sample sizes or appropriate analytic methods. Selection bias and failure to isolate variables have further plagued these studies, as have inadequate or imprecise measures of students' mathematics achievement. Finally, no studies identified by the Task Group probed the dynamic that would examine how teachers' mathematical knowledge affects instructional quality, students' opportunities to learn, and their gains over time. However, in the context of a body of literature that is as inexact as it has been, the positive trends identified do support the importance of teachers' knowledge of mathematics as a factor in students' achievement.

To improve the quality of research evidence, sharper measures of teachers' mathematical knowledge in different domains and at different levels are needed, with appropriate psychometric tests for their reliability and validity. Hill et al. (in press) report a series of in-depth validation studies that provide a benchmark for what such tests should involve. Also needed are value-added studies, including experimental studies of interventions designed to develop teachers' mathematical knowledge for teaching and to inspect the effects of interventions on students' mathematics achievement.

As the scientific rigor and validity of research in this area increases, it should be informed by analyses of empirical data on the mathematical demands on teachers in the course of their work. Hypotheses generated through a mathematical perspective on teaching practice can provide researchers with a sharper focus on the question in ways more likely to produce results useful to policymakers, and to improve the relevance and effectiveness of teachers' mathematical training and professional assessment.

6. Recommendations Based on the Mathematical and Logical Analysis of The Demands of Teaching Mathematics

Although the empirical evidence does not yet strongly specify the nature of the mathematical knowledge needed for teaching, the Task Group hypothesizes that teaching mathematics demands knowledge of the subject. Because a direct relationship between conventional mathematical study and teacher effectiveness is not supported by a review of high-quality research, future research should uncover those aspects of teacher knowledge and understanding that are most strongly related to student learning. Policies should be developed and supported that would lead to a more mathematically skilled teacher force. The Task Group recommends investigating different components of content knowledge that may have relationships to instructional effectiveness.

A first component worth investigating is the nature of the requisite competence with the school curriculum that teachers are responsible to teach. This includes the concepts, skills, and strategies that students are to learn, and the prerequisite content, several levels both below and beyond the level at which they teach. Another aspect of this curricular knowledge is the additional need for teachers to know school mathematics at a more advanced level than what is found in school textbooks. This may be seen in the daily tasks that teachers must perform and that appear to entail substantial mathematical judgment, understanding, and skill. For example, answering students' questions may be unexpectedly subtle or complicated, or making up exam problems to focus on central ideas may require more than a minimal knowledge of the mathematics in the grade-level curriculum. It is also worth examining the extent to which teachers may need different types of advanced knowledge—for example, how an understanding of number theory may enrich and strengthen teachers' capacity to teach whole number concepts and operations.

Another important area in need of investigation is what else teachers may need beyond the standard skills and concepts in the school curriculum. Examples of work that teachers do that may require other mathematical knowledge include explaining why a particular topic is worth learning; providing connections within lessons, across lessons, and across grades; and making spontaneous instructional decisions in the classroom. These tasks of teaching seem to require mathematical skill that has not been well established in the research literature.

To help students extend what they know, teachers may also need a deep understanding of the foundational ideas and skills prerequisite to the level of instruction, the mathematics that leads up to the students' mathematical present. Middle school teachers teaching algebra may benefit from a nuanced understanding of operations, including their properties, and their interpretation and representations; upper elementary teachers may be enabled by an understanding of ways of renaming or re-representing mathematical ideas that arise in the early grades (e.g., with the standard subtraction algorithm).

In addition to knowledge of particular content, including concepts and procedures, teaching may also require sensitivity, habits of mind, and attention to particular mathematical principles, such as precision, definitions, reasoning, and coherence. Mathematics uses language in exact ways quite different from its uses in other contexts, including everyday life. Attention to careful use of quantifiers, relations, and logical terms is important. Being clear on the differences between "five apples," "at least five apples," "no more than five apples," and "exactly five apples," or between "if" and "if and only if" is important as young learners begin to express mathematical ideas, or as a teacher attentively uses a curriculum. A second, and closely related, aspect of mathematics is the importance of *definitions*. Terms, ideas, and concepts all require definition in the service of precision and to support reasoning.

The Task Group reviewed how teachers' knowledge of particular definitions, and perhaps equivalent alternatives and their relative advantages, make a difference for teaching quality. Another aspect for consideration may be a sense about the general role and nature of definitions within mathematics. Mathematical development and the solution of problems depend on specific approaches to logical argument and explanation. How does teachers' knowledge of reasoning play a role in teaching, considering both the teachers' own knowledge and their capacity to make mathematical reasoning and explanation accessible to and learnable by students? Being able to do that is rooted in a finely grained understanding of the nature of mathematical reasoning. Also worth investigating is the ability to see connections, and to appreciate and construct coherent links within and across ideas. How is early work with whole numbers connected to later encounters with integers, fractions, and the real number line? How are equivalent fractions related to the regrouping steps of the subtraction algorithm? What are useful geometric or physical models of arithmetic operations, and how can one explain the correspondences?

More work is needed to inspect how mathematical knowledge is needed for and deployed in teaching. Hypotheses about the requisite skill and knowledge can help to advance the question from a blunt investigation of credentials to a more nuanced understanding of the mathematical demands of teaching and the connection of those to students' learning gains. This is a crucial area for further and more precise study.

B. Teachers' Education: Teacher Preparation and Alternative Pathways to Teaching, Professional Development, and Induction

What kinds of programs have been shown to help teachers develop the necessary mathematical knowledge and skills needed for teaching?

a) How can preservice programs effectively increase beginning teachers' mathematical knowledge for teaching?
b) How can in-service programs do so?
c) Do particular designs or curricula make a difference for teachers' instructional skill and their students' achievement?
d) Is there evidence about how different kinds of professional preparation or requirements affect teachers' effectiveness, and how these compare?

Teacher education is regarded as key to building instructional quality and teacher effectiveness. The Task Group uses the term "teacher education" here to refer to four different types of professional training:

- *Preservice teacher preparation*: Initial teacher training, conventionally offered in institutions of higher education;

- *Alternative pathways*: Initial teacher preparation, offered outside of conventional teacher education programs;

- *Induction programs*: Programs of professional support and additional training within the first years of practice; and

- *Professional development*: Ongoing programmatic professional education of practicing teachers.

The Task Group sought to examine the evidence on teacher education in these four forms, asking about the relationship between different forms of teacher education and the learning of teachers and their students.

Many beliefs exist about what constitutes effective professional training for teachers, including what teachers should learn, how it should be structured and taught, and how much training is needed. There are also beliefs about learning from experience, and about what can be learned in formal settings or from practice. Many authoritatively stated positions assert "what we know" about "good" professional development. The Task Group wanted to learn what is known about particular curricula, structures, or approaches to teacher education, and their effects on gains in teachers' mathematical knowledge and skill for teaching, and their demonstrated relationships to students' achievement gains. The Task Group focused on these two outcomes to learn about effective professional training; although other outcomes (e.g., professional satisfaction, retention, teachers' reports of usefulness or relevance of particular programs) may be informative, they fall short of providing links between professional education and actual effects on learning outcomes. To help inform sensible allocation of resources for professional education and to provide direction for continued research in this area, the Task Group focused directly on the current state of knowledge about these effects of professional education. The Task Group's results highlight the critical need for more and better studies tracing the relationship between specific approaches to teacher education (i.e., curricula, pedagogy and assessment, instructors, structures, and settings) and teachers' capacity for teaching and their students' learning.

1. Preservice Teacher Preparation

The Task Group identified and synthesized peer-reviewed research and national reports to answer questions regarding the impact of preparation programs for teachers. Most of the studies found were descriptive—that is, they provided information about programs, described the characteristics of individuals who enrolled in or completed such programs, or simply compared students before and after a program or class without any comparison group. Other studies looked at limited relationships, such as the effect of being taught by a teacher who is certified in mathematics. Such studies were not useful for the question the Task Group sought to answer about features of preservice teacher preparation and their effects on teacher knowledge or student achievement.

Five empirical studies were found that addressed the question related to impacts of preparation programs on student achievement or teachers' mathematical content knowledge. Two of the studies examined effects on student achievement (Levine, 2006; Noell, 2006); the other three examined impacts on teacher mathematics content knowledge (Koehler & Lehrer, 1998; McDevitt, Troyer, Ambrosio, Heikkinen, & Warren, 1995; National Center for Research on Teacher Learning, 1991). These studies employed a variety of analytical techniques, but none was of sufficient rigor or quality to allow the Task Group to draw conclusions about the relationship of particular features of teacher preparation programs and their effects. Only two of the five were peer-reviewed and none controlled for all the relevant factors that might explain variation in impact on either teachers' knowledge or their students' learning.

In Tables 7 and 8, a synthesis is provided of the empirical evidence uncovered providing information on both study characteristics and study findings. Overall, however, the Task Group was unable to draw conclusions from this body of evidence.

Table 7: Quality Characteristics of Models Examining Impact of Teacher Preparation Programs on Student Achievement in Mathematics or Teacher Mathematics Content Knowledge

			Other Controls					
Authors	**Sample Size**	**Pretest Control**	**Student**	**Teacher**	**Class, School, or District**	**Family**	**Matching of Schools**	**Level of Analysis**
Student Achievement in Mathematics								
Levine, 2006	1,611 math teachers, over 30 million observations of students	X		X	X	X	X	Teacher
Noell, 2006	Over 200,000 students in Grades 4–9	X	X		X	X		Student
Teacher Mathematics Content Knowledge								
Koehler, & Lehrer, 1998	10 preservice teachers	X						Preservice teacher
McDevitt et al., 1995	About 150 preservice teachers in the first cohort, 110 in the second cohort							Preservice teacher
National Center for Research on Teacher Learning, 1991	Longitudinal study of over 100 participants in 5 different programs	X						Preservice teacher

Table 8: Reported Impacts From Models Examining the Effect of Teacher Preparation Programs on Student Achievement in Mathematics or Teacher Mathematics Content Knowledge

Authors	Dependent Variable	Independent Variable	Results of Estimated Effects	Math Specific Outcome	Grade(s)	Estimation Technique
Student Achievement in Mathematics						
Levine, 2006	NWEA achievement tests	Teacher trained at NCATE accredited school	Positive & Not Significant	X		ANOVA[a]
	NWEA achievement tests	Teacher trained at doctoral/research university (vs. Masters I)	Positive & Significant	X		ANOVA
Noell, 2006	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ B	Positive & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ E	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ H	Positive & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ K	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ L	Positive & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ M	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ N	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Alternative Cert. Univ M	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Alternative Cert. Univ P	Positive & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ A	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ D	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ F	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ G	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ J	Negative & Not Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Undergraduate Univ I	Negative & Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Alternative Cert. Univ B	Negative & Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Alternative Cert. Univ G	Negative & Significant	X	4–9	Hierarchical Linear Models (HLM)
	Louisiana standardized test scores in math	Teacher attended Alternative Cert. Univ L	Negative & Significant	X	4–9	Hierarchical Linear Models (HLM)
Teacher Mathematics Content Knowledge						
Koehler, & Lehrer, 1998	Problem-type sorting task	Learning using a hypermedia tool (vs. text)	Positive & Not Significant	X	N/A[b]	Comparison of Means
	Solution-strategy sorting task	Learning using a hypermedia tool (vs. text)	Positive & Significant	X	N/A	Comparison of Means
McDevitt et al., 1995	Researcher-designed test given at the end of each math class	An experimental program (vs. standard courses)	Positive & Significant	X	N/A	Comparison of Means
National Center for Research on Teacher Learning, 1991	Score on interview tasks and questionnaire	Attended the only program in which developing pre-service teachers' meaningful knowledge was an explicit goal	Positive, Significance not reported	X	N/A	Comparison of Means

[a] Analysis of variance.
[b] N/A means that data were not available.

2. Alternative Pathways Into Teaching

The Task Group identified peer-reviewed research and national reports focused on the impact of alternative preparation programs for teachers. However, most of the studies located were only descriptive. These reports provided information about alternative programs or described the characteristics of individuals who enrolled in or completed such programs.

The Task Group found 10 empirical studies that examined effects of alternative preparation programs on student achievement; all but one used correlational techniques (one was experimental). No empirical studies were found on the impacts of alternative pathways programs on teachers' mathematical content knowledge. Tables 9 and 10 provide a synthesis of the relevant empirical evidence, and information on both study characteristics and study findings. Wherever possible, standardized regression coefficients were calculated. Overall, evidence is mixed on programmatic effects of teachers' pathways into teaching and their relationship to students' achievement. As shown in Table 10, four studies show positive effects for teachers trained in alternative pathways [e.g., Teach for America (TFA)]. Boyd, Grossman, Lankford, Loeb, & Wykoff (2006) found small differences in students' mathematics achievement that could be associated with teacher preparation pathways and these effects were only for first-year teachers working with students in the sixth through eighth grades. Out of 18 different comparisons of different alternative pathways and "college-recommended pathways" investigated, 5 showed significant effects, with 4 of those showing positive effects for an alternative pathway, and 1 showing a negative effect. Overall, according to the authors, variation within pathways tended to be greater than variation across pathways. Decker, Mayer, and Glazerman (2004) found that students in TFA classrooms outperformed their peers, though the size of these effects varied with the characteristics of students (e.g., gender, mobility, and prior achievement status).

Kane, Rockoff, and Staiger (2006) showed significant positive effects for one alternative pathway into teaching (Teach for America, TFA) and significant negative effects for a second alternative pathway (International Program). Results were also inconclusive in the research reported by Darling-Hammond, Holtzman, Gatlin, and Heilig (2005), who found more negative effects for teachers prepared through alternative pathways than for those prepared in traditional certification programs. In Raymond, Fletcher, and Luque (2001), the students of new TFA teachers outperformed those of new traditionally trained teachers at Grades 4 and 5, while there was no difference for teachers overall. At the middle school level, however, the difference was significant for all teachers overall but insignificant for new teachers. In contrast, Miller, McKenna, and McKenna (1998) found no differences between the mathematics achievements of fourth- or fifth-grade students whose teachers were prepared through alternative pathways and those prepared in traditional programs. Additionally, Laczko-Kerr and Berliner (2002) found that students of under-qualified teachers, including TFA teachers, performed less well on mathematics tests than those of comparably experienced certified teachers.

Interpreting the evidence across these studies is not easy. One significant problem is definitional and related to the treatment conditions; "alternative pathways" does not define a clear programmatic type of teacher preparation. Some studies examined certification status while others compared different pathways to certification. Similarly, there is lack of clear specification of the traditional preparation programs to which these alternatives are being compared. These programs—alternative and traditional—are all forms of preservice teacher preparation, and the studies do not probe into key curricular, structural, or other programmatic variables that would permit analysis of the programs and their effects. A second problem rests with massive differences in the design and measures of students' achievement, making it difficult to compare across studies. A third issue is that the programs that tend to be studied over-sampled from one state (New York) and, thus, represent a narrow range geographically and only a tiny fraction of teachers in the system with over-sampling from more elite populations. Drawing conclusions about alternative pathways, in general, would be difficult to do from these studies.

Determining how different types of pathways into teacher preparation may affect teachers' capacity to teach is a key policy question in a time when issues of teacher recruitment, retention, and quality are paramount. Extant evidence suggests that there are no significant differences among current pathways, and, as Boyd et al. (2006) report, variation within programs appears to be greater than that uncovered across programs. Studies that more clearly specified the alternatives to be compared and the outcome measures used, and that used common or comparable designs would help in investigating this important question with more precision and focus.

Table 9: Quality Characteristics of Models Examining Impact of Alternative Pathways On Student Achievement in Mathematics

Authors	Sample Size	Pretest Control	Other Controls: Student	Other Controls: Teacher	Other Controls: Class, School, or District	Other Controls: Family	Matching of Teachers	Level of Analysis
Boyd, Grossman, Loeb, Lankford, & Wyckoff, 2006	1,035,949 Grades 3–8 student-year observations	Yes	X	X	X	X		Student
Fetler, 1999	Total: 921,437 Grade 9: 347,201 Grade 10: 313,303 Grade 11: 260,933 students in schools across CA			X	X	X		School
Goldhaber, & Brewer, 2000	3,786 Grade 12 math students 2,098 math teachers	X	X	X	X	X		Student
Kane, Rockoff, & Staiger, 2006	1,462,100 student-year observations, for Grades 3–8 in NYC schools from 1998–2005	X	X	X	X			Student
Tatto, Nielsen, Cummings, Kularatna, & Dharmadasa, 1993	216 teachers in Sri Lanka		X	X	X			Student
Decker, Mayer, & Glazerman, 2004	1,893 students in across 100 classrooms	X	X	X	X	X		Student
Darling-Hammond, Holtzman, Gatlin, & Heilig, 2005	271,015 students in the Houston ISD, Grades 3 and higher from 1995–96 to 2001–02	X	X	X	X	X		Student
Laczko-Kerr, & Berliner, 2002	232 newly hired teachers across 5 AZ school districts from 1998–2000			X	X		X	Classroom
Raymond, Fletcher, & Luque, 2001	81,814 students in Grades 4 & 5 96,276 students in Grades 6 & 8	X	X	X	X	X		Student
Miller, McKenna & McKenna, 1998	345 students in 18 middle school classrooms in GA			X			X	Student

Table 10: Reported Impacts of Studies Examining the Effect of Teachers From Alternative Paths on Student Achievement in Mathematics

Authors	Dependent Variable	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade(s)	Estimation Technique
Boyd, Grossman, Loeb, Lankford, & Wyckoff, 2006	NYC standardized math exam scores	NYC Teaching Fellow (vs. College Recomm.) Over 1 year	Unknown	Negative & Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	NYC Teaching Fellow (vs. College Recomm.) Over 2 years	Unknown	Positive & Not Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	NYC Teaching Fellow (vs. College Recomm.) Over 3 years	Unknown	Positive & Not Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	TFA Teacher (vs. College Recomm.) Over 1 year	Unknown	Negative & Not Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	TFA Teacher (vs. College Recomm.) Over 2 year	Unknown	Positive & Not Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	TFA Teacher (vs. College Recomm.) Over 3 years	Unknown	Positive & Not Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	Temp. Licensed (vs. College Recomm.) Over 1 year	Unknown	Negative & Not Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	Temp. Licensed (vs. College Recomm.) Over 2 years	Unknown	Positive & Not Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	Temp. Licensed (vs. College Recomm.) Over 3 years	Unknown	Positive & Significant	X	4–5	OLS Regression
	NYC standardized math exam scores	NYC Teaching Fellow (vs. College Recomm.) Over 1 year	Unknown	Negative & Not Significant	X	6–8	OLS Regression
	NYC standardized math exam scores	NYC Teaching Fellow (vs. College Recomm.) Over 2 years	Unknown	Positive & Not Significant	X	6–8	OLS Regression
	NYC standardized math exam scores	NYC Teaching Fellow (vs. College Recomm.) Over 3 years	Unknown	Positive & Significant	X	6–8	OLS Regression
	NYC standardized math exam scores	TFA Teacher (vs. College Recomm.) Over 1 year	Unknown	Positive & Significant	X	6–8	OLS Regression
	NYC standardized math exam scores	TFA Teacher (vs. College Recomm.) Over 2 years	Unknown	Positive & Not Significant	X	6–8	OLS Regression

Continued on p. 5-31

Table 10, continued

Authors	Dependent Variable	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade(s)	Estimation Technique
Boyd, Grossman, Loeb, Lankford, & Wyckoff, 2006	NYC standardized math exam scores	TFA Teacher (vs. College Recomm.) Over 3 years	Unknown	Positive & Not Significant	X	6–8	OLS Regression
	NYC standardized math exam scores	Temp. Licensed (vs. College Recomm.) Over 1 year	Unknown	Negative & Not Significant	X	6–8	OLS Regression
	NYC standardized math exam scores	Temp. Licensed (vs. College Recomm.) Over 2 years	Unknown	Positive & Significant	X	6–8	OLS Regression
	NYC standardized math exam scores	Temp. Licensed (vs. College Recomm.) Over 3 years	Unknown	Positive & Not Significant	X	6–8	OLS Regression
Fetler, 1999	Stanford Achievement Test scores in math	Percent Emergency Certified Teachers	Unknown	Negative & Indeterminate	X	9	OLS Regression
	Stanford Achievement Test scores in math	Percent Emergency Certified Teachers	Unknown	Negative & Indeterminate	X	10	OLS Regression
	Stanford Achievement Test scores in math	Percent Emergency Certified Teachers	Unknown	Negative & Indeterminate	X	11	OLS Regression
Goldhaber, & Brewer, 2000	Standardized test scores in math, for test designed by the ETS (NELS:88)	Probationary Certification (vs. Traditionally Certified)	0.00	Positive & Not Significant	X	12	OLS Regression
	Standardized test scores in math, for test designed by the ETS (NELS:88)	Emergency Certification (vs. Traditionally Certified)	0.00	Positive & Not Significant	X	12	OLS Regression
	Standardized test scores in math, for test designed by the ETS (NELS:88)	Not Certified (vs. Traditionally Certified)	-0.01	Negative & Significant	X	12	OLS Regression

Continued on p. 5-32

Table 10, continued

Authors	Dependent Variable	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade(s)	Estimation Technique
Kane, Rockoff, & Staiger, 2006	NYC standardized math exam scores	NYC Teaching Fellow (vs. Traditionally Certified)	0.00	Positive & Not Significant	X	3–8	OLS Regression
	NYC standardized math exam scores	TFA Teacher (vs. Traditionally Certified)	0.01	Positive & Significant	X	3–8	OLS Regression
	NYC standardized math exam scores	Int'l Program Teacher (vs. Traditionally Certified)	0.00	Negative & Significant	X	3–8	OLS Regression
	NYC standardized math exam scores	Not Certified (vs. Traditionally Certified)	0.00	Zero	X	3–8	OLS Regression
Tatto, Nielsen, Cummings, Kularatna, & Dharmadasa, 1993	Sri-Lankan test of math skills	Untrained Teacher (vs. Teacher College)	Unknown	Positive & Significant	X	4	Comparison of Means
	Sri-Lankan test of math skills	Untrained Teacher (vs. College of Education)	Unknown	Positive & Significant	X	4	Comparison of Means
	Sri-Lankan test of math skills	Untrained Teacher (vs. Distance Education)	Unknown	Positive & Not Significant	X	4	Comparison of Means
Decker, Mayer, & Glazerman, 2004	ITBS math scores	TFA Teacher	Unknown	Positive & Significant	X	1–5	HLM Modeling
Darling-Hammond, Holtzman, Gatlin, & Heilig, 2005	TAAS score in math	TFA Teacher	Unknown	Positive & Significant	X	3–HS	HLM Modeling
	Stanford Achievement Test score in math	TFA Teacher	Unknown	Negative & Significant	X	3–HS	HLM Modeling
	Aprenda score in math	TFA Teacher	Unknown	Negative & Significant	X	3–HS	HLM Modeling
	TAAS score in math	Alternative Certification	Unknown	Negative & Significant	X	3–HS	HLM Modeling
	Stanford Achievement Test score in math	Alternative Certification	Unknown	Negative & Significant	X	3–HS	HLM Modeling
	Aprenda score in math	Alternative Certification	Unknown	Negative & Not Significant	X	3–HS	HLM Modeling
	TAAS score in math	Emer./Temp. Certification	Unknown	Negative & Significant	X	3–HS	HLM Modeling
	Stanford Achievement Test score in math	Emer./Temp. Certification	Unknown	Negative & Not Significant	X	3–HS	HLM Modeling
	Aprenda score in math	Emer./Temp. Certification	Unknown	Negative & Not Significant	X	3–HS	HLM Modeling

Continued on p. 5-33

Table 10, continued

Authors	Dependent Variable	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade(s)	Estimation Technique
Laczko-Kerr, & Berliner, 2002	Average Stanford Achievement Test score in math in 1998	Certified Teacher (vs. Noncertified)	Unknown	Positive & Not Significant	X	2–8	Analysis of variance (ANOVA) and Comparisons of means
	Average Stanford Achievement Test score in math in 1999	Certified Teacher (vs. Noncertified)	Unknown	Positive & Significant	X	2–8	ANOVA and Comparisons of means
	Average Stanford Achievement Test score in math in 1998	Not TFA (vs. TFA teachers)	Unknown	Positive & Not Significant	X	2–8	ANOVA and Comparisons of means
	Average Stanford Achievement Test score in math in 1999	Not TFA (vs. TFA teachers)	Unknown	Positive & Significant	X	2–8	ANOVA and Comparisons of means
Raymond, Fletcher, & Luque, 2001	Average TAAS math score	TFA Teacher (vs. non-TFA teachers, for all teachers)	Unknown	Positive & Not Significant	X	4–5	OLS regression
	Average TAAS math score	TFA Teacher (vs. non-TFA teachers, w/ < 1 yr exp)	Unknown	Positive & Significant	X	4–5	OLS regression
	Average TAAS math score	TFA Teacher (vs. non-TFA teachers, for all teachers)	Unknown	Positive & Significant	X	6–7	OLS regression
	Average TAAS math score	TFA Teacher (vs. non-TFA teachers, w/ < 1 yr exp)	Unknown	Positive & Not Significant	X	6–7	OLS regression
Miller, McKenna, & McKenna, 1998	ITBS reading and math scores	Alternative Certification	Unknown	Positive & Not Significant	X	5–6	Multivariate Analysis of variance (MANOVA) and Comparisons of means

3. Induction Programs

National reports calling for higher-quality teaching, higher teacher retention rates, and stronger student achievement identify support of new teachers—or induction—as an area for improvement (National Commission on Mathematics and Science Teaching for the 21st Century, 2000; National Commission on Teaching and America's Future, 2003; National Council of Teachers of Mathematics, 2002). As the goal of many induction programs is acclimation of new teachers to the school, they often last only the first year. Other induction programs are positioned as the first step on a continuum of professional development for teachers and are multiyear in nature. The assignment of mentors is frequently part, or all, of an induction program.

Feiman-Nemser, Schwille, Carver, and Yusko (1999) discuss three definitions for induction: 1) a unique phase or stage in teacher development, 2) a time for socialization as the teacher transitions from preparation to practice, and 3) a formal program for beginning teachers. In this review, as is common, the term “induction” was taken to mean a formal program for beginning teachers, including mentoring programs. Induction programs are varied in the length of time a person participates, the scope of support provided, and the guidance provided by policies and mandates (Feiman-Nemser et al.). Some states, such as California and Connecticut, have statewide induction programs.

Frequently, at least in the United States, induction is viewed as synonymous with mentoring; however, this is not accurate. Mentoring is a frequent component of induction programs, but it is not the only component. Four goals of induction programs were identified by the National Commission on Teaching and America’s Future (Fulton, Yoon, & Lee, 2005): 1) building and deepening teacher knowledge, 2) integration of new practitioners into a teaching community and school culture that support continuous professional growth of all, 3) support for the constant development of the teaching community in the school, and 4) encouragement of a professional dialogue to articulate the goals, values, and best practices of the community. Induction programs should be systems which are “networks of supports, people, and processes that are all focused on assuring that novices become effective in their work” (p. 4).

The concept of induction is not new in education. Although calls for programs to support new teachers can be dated to the 1960s, initially induction programs were uncommon. Only Florida had a mandated induction program prior to 1980 (Feiman-Nemser et al., 1999). Over the years induction programs, and mandates to support induction programs, increased due to ties between induction and such key issues as school reform, teacher retention, teacher quality, and achievement initiatives.

The Task Group searched for empirical investigations of the effectiveness of teacher induction programs—broadly defined—on teacher mathematics knowledge and student achievement. Careful systematic review of the literature uncovered a dearth of peer-reviewed research on induction. Reviews of induction programs—frequently focusing on retention—were examined (e.g., Feiman-Nemser et al., 1999; Glazerman, Senesky, Seftor, & Johnson, 2006; Ingersoll & Kralik, 2004; Kagan, 1992; Lopez, Lash, Schaffner, Shields, & Wagner, 2004; Totterdell, Bubb, Woodroffe, & Hanrahan, 2004). Overall, they identified more than 100 potentially relevant pieces of literature. However, none focused on the effects of induction for mathematics teachers on student achievement or teacher mathematics knowledge.

Literature examining induction is not scarce. However, rarely does it focus specifically on teachers of mathematics—although some focuses on both mathematics and science teachers (Adams & Krockover, 1997; Davis, Petish, & Smithey, 2006; Garet, Porter, Desimone, Birman, & Yoon, 2001; McGinnis, Parker, & Graeber, 2004). Much of the work is program evaluation of particular initiatives and is not peer-reviewed. For examples, see two programs in California: the New Teacher Center at the University of California Santa Cruz (www.newteachercenter.org), and California’s Beginning Teacher Support and Assessment program (www.btsa.ca.gov). In addition, many of the studies are case studies or qualitative in nature. The empirical evidence on outcomes of the programs is weak as none of the studies reviewed used random assignment, and few used a comparison group of any kind.

The key outcome for much of the extant induction literature is teacher retention. There is also a wealth of literature examining the effects of induction programs on teacher beliefs, satisfaction, and practices. Induction programs continue to expand, some mandated and some not. Given the expansion, it is important to assess the effectiveness of induction programs on outcomes, such as student achievement, and not rely only on evidence about teacher retention, satisfaction, and beliefs. Until induction programs are content-specific or include specific content activities, it may be difficult to determine the effectiveness of induction on the mathematical knowledge of teachers or on students' achievement gains.

4. Professional Development Programs

The final component of teacher education that the Task Group investigated was professional development for practicing teachers. The Task Group searched for peer-reviewed research and national reports that would offer high-quality evidence regarding the impact of professional development programs for teachers. Many of the studies identified were descriptive in that they provided information about the characteristics of the programs, and most of those that were empirical did not include a comparison group, but used a one-group pretest-posttest design. This was the case for every empirical study they identified that examined the effects of teacher professional development programs on teachers' mathematical content knowledge. Moreover, many of those relied on teacher self-reports about their knowledge before and after the professional development rather than on measures of teacher knowledge. As a result, this review includes only studies investigating the relationship between teacher professional development programs and students' mathematics achievement. In other words, the Task Group did not include studies with a pretest-posttest design, and thus no studies related to teacher mathematics content knowledge were included.

The Task Group found eight empirical studies, using a variety of analytical techniques, that examined effects of teacher professional development programs on student achievement. Tables 11 and 12 provide a synthesis of the relevant empirical evidence, and information on both study characteristics and study findings. For all studies except one (Jacob & Lefgren, 2002), the Task Group was able to calculate a standardized effect size using Hedge's *g*. These calculations are shown in Table 12. Accompanying this synthesis is a series of descriptive tables that provide more details on each study included in Tables 11 and 12.

Table 11: Quality Characteristics of Studies Examining Impact of Teacher Professional Development Programs on Student Achievement in Mathematics

			Other Controls					
Authors	**Sample Size**	**Pretest Control**	**Student**	**Teacher**	**Class, School, or District**	**Family**	**Matching of Schools**	**Level of Analysis**
Angrist, & Lavy, 2001	Elementary schools in Jerusalem, 9 intervention schools (7 secular, 2 religious) and 11 comparison schools (6 secular, 5 religious). Approximately 634 secular students; 196 religious students.	X	X		X	X	X	Student
Campbell, 1996	Elementary school teachers (K–Grade 3) in 6 schools in Montgomery County, MD. Treatment group: students in 3 treatment schools; comparison group: student in 3 schools in the same county. *n* = 149 Kindergarten and 292 Grade 1 students.						X	Student
Carpenter et al., 1989	40 Grade 1 teachers in 24 schools (2 private) in Madison, WI.	X						Teacher
Chapin, 1994	Elementary, middle and high school students of 42 teachers in Chelsea, MA, school district versus students in the same schools whose teachers were not involved in the program. Outcomes available for Grades 3 (*n* = 269), 6 (*n* = 244) and 7 (*n* = 210) students.							Student
Jacob, & Lefgren, 2002	246 Elementary schools in Chicago (Grades 3–6), approximately 47,000 students.	X	X		X	X	Regular discontinuity	Student
Karges-Bone, Collins, & Maness, 2002	1 elementary school, 61 Grade 3 students, 48 Grade 4 students.							Student
Saxe, Gearheart, & Nasir, 2001	23 elementary schools in Greater Los Angeles: 9 teachers in Integrated Mathematics Assessment (IMA) program, 8 teachers in collegial support program (SUPP); and 6 in traditional/no additional professional development program (TRAD).	X	X					Classroom
Van Haneghan, Pruet, & Bamberger, 2004	6 elementary schools (4 treatment, 2 comparison) in Mobile, AL. Followed two cohorts over the course of 3 years (approx 200 in K, and approx. 420 in Grade 3).						X	Student

Table 12: Reported Impacts of Studies Examining the Effect of Teacher Professional Development Programs on Student Achievement in Mathematics

Authors	Dependent Variable	Independent Variable	Effect size or Regression Coefficients	Results of Estimated Effects	Math Specific Outcomes	Grade(s)	Estimation Technique
Angrist & Lavy, 2001	1996 math score	Attending secular school with more intensive in-service training	0.431	Positive (p = .097)	X	6	Hedges g based on raw means (adjusted for differences in average pretest scores and adjusted for school-level clustering)
Campbell, 1996	Project-developed student achievement assessment	Project IMPACT vs. comparison	0.39	Positive, not significant	X	K	Hedges g (adjusted for school clustering)
	Project-developed student achievement assessment	Project IMPACT vs. comparison	0.14	Positive, not significant	X	1	Hedges g (adjusted for school clustering)
Carpenter et al., 1989	Computation – ITBS	CGI treatment vs. basic problem-solving workshop	0.41	Positive, not significant	X	1	Hedges g (pretest adjusted means, no adjust for clustering)
	Computation – Number facts	CGI treatment vs. basic problem-solving workshop	0.66	Positive, significant	X	1	Hedges g (pretest adjusted means, no adjust for clustering)
	Problem solving: ITBS	CGI treatment vs. basic problem-solving workshop	0.37	Positive, not significant	X	1	Hedges g (pretest adjusted means, no adjust for clustering)
	Problem solving: simple add/subtract	CGI treatment vs. basic problem-solving workshop	0.43	Positive, not significant	X	1	Hedges g (raw means since adjusted not available, no adjust for clustering)
	Problem solving: complex add/subtract	CGI treatment vs. basic problem-solving workshop	0.42	Positive, not significant	X	1	Hedges g (pretest adjusted means, no adjust for clustering)
	Problem solving: advanced	CGI treatment vs. basic problem-solving workshop	0.11	Positive, not significant	X	1	Hedges g (pretest adjusted means, no adjust for clustering)
	Problem solving: interview	CGI treatment vs. basic problem-solving workshop	0.69	Positive, significant	X	1	Hedges g (pretest adjusted means, no adjust for clustering)
Chapin, 1994	California Achievement Test	Students of CCTDM teachers versus non-project teachers	0.58	Positive, significant	X	3	Hedges g (adjusted for classroom clustering)
	California Achievement Test	Students of CCTDM teachers versus non-project teachers	0.76	Positive, significant	X	6	Hedges g (adjusted for classroom clustering)
	California Achievement Test	Students of CCTDM teachers versus non-project teachers	0.66	Positive, significant	X	7	Hedges g (adjusted for classroom clustering)
Jacob, & Lefgren, 2002	ITBS math score	Probation vs. no	OLS est = -.021 (se = .010)	Negative, not significant	X	3–6	OLS with student and school covariates
Karges-Bone, Collins, & Maness, 2002	BSAP	1997/98 cohort (post-treatment) vs. 1996 cohort	0.70	Positive, significant	X	3	Hedges g (student level)
	MAT	1997/98 cohort (post-treatment) vs. 1996 cohort	0.44	Positive, not significant	X	4	Hedges g (student level)

Continued on p. 5-38

Table 12, continued

Authors	Dependent Variable	Independent Variable	Effect size or Regression Coefficients	Results of Estimated Effects	Math Specific Outcomes	Grade(s)	Estimation Technique
Saxe, Gearheart, & Nasir, 2001	Concepts	IMA vs. TRAD	2.39	Positive, significant	X	4–5	Hedges g (classroom-level analysis)
	Concepts	SUPP vs. TRAD	0.67	Positive, not significant	X	4–5	Hedges g (classroom-level analysis)
	Concepts	IMA vs. SUPP	1.45	Positive, significant	X	4–5	Hedges g (classroom-level analysis)
	Computation	IMA vs. TRAD	-0.53	Negative, not significant	X	4–5	Hedges g (classroom-level analysis)
	Computation	SUPP vs. TRAD	-1.34	Negative, significant	X	4–5	Hedges g (classroom-level analysis)
	Computation	IMA vs. SUPP	0.77	Positive, not significant	X	4–5	Hedges g (classroom-level analysis)
Van Haneghan, Pruet, & Bamberger, 2004	Yr 1, Grade 3: SAT9 Problem Solving NCE	Students of trained MMI teachers vs. control schools	-0.04	Negative, not significant	X	3	Hedges g (student level, adjusted for school-level clustering)
	Yr 1, Grade 3: SAT9 Procedures NCE	Students of trained MMI teachers vs. control schools	0.29	Positive, not significant	X	3	Hedges g (student level, adjusted for school-level clustering)
	Yr 1, Grade 3: SAT9 Total NCE	Students of trained MMI teachers vs. control schools	0.15	Positive, not significant	X	3	Hedges g (student level, adjusted for school-level clustering)
	Yr 1, Grade 3: TIMSS-items	Students of trained MMI teachers vs. control schools	0.35	Positive, not significant	X	3	Hedges g (student level, adjusted for school-level clustering)
	Yr 2, Grade 4: SAT9 Problem Solving NCE	Students in MMI schools versus control schools	0.24	Positive, not significant	X	4	Hedges g (student-level, adjusted for school-level clustering)
	Yr 2, Grade 4: SAT9 Procedures NCE	Students in MMI schools versus control schools	0.24	Positive, not significant	X	4	Hedges g (student level, adjusted for school-level clustering)
	Yr 2, Grade 4: SAT9 Total NCE	Students in MMI schools versus control schools	0.27	Positive, not significant	X	4	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 2: Fractions	Students in MMI schools versus control schools	0.20	Positive, not significant	X	2	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: Fractions	Students in MMI schools versus control schools	0.70	Positive, not significant	X	5	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 2: Geometry	Students in MMI schools versus control schools	0.01	Positive, not significant	X	2	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 2: Mental Math	Students in MMI schools versus control schools	0.22	Positive, not significant	X	2	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: Mental Math	Students in MMI schools versus control schools	0.46	Positive, not significant	X	5	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 2: Numeration	Students in MMI schools versus control schools	0.11	Positive, not significant	X	2	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: Numeration	Students in MMI schools versus control schools	0.36	Positive, not significant	X	5	Hedges g (student level, adjusted for school-level clustering)

Continued on p. 5-39

Table 12, continued

Authors	Dependent Variable	Independent Variable	Effect size or Regression Coefficients	Results of Estimated Effects	Math Specific Outcomes	Grade(s)	Estimation Technique
Van Haneghan, Pruet, & Bamberger, 2004	Yr 3, Grade 2: Story Problems	Students in MMI schools versus control schools	0.27	Positive, not significant	X	2	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: Story Problems	Students in MMI schools versus control schools	0.47	Positive, not significant	X	5	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: SAT9 Problem Solving NCE	Students in MMI schools versus control schools	0.57	Positive, not significant	X	5	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: SAT9 Procedures NCE	Students in MMI schools versus control schools	0.67	Positive, not significant ($p = .097$)	X	5	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: SAT9 Total NCE	Students in MMI schools versus control schools	0.66	Positive, not significant	X	5	Hedges g (student level, adjusted for school-level clustering)
	Yr 3, Grade 5: TIMSS items	Students in MMI schools versus control schools	0.71	Positive, not significant ($p = .078$)	X	5	Hedges g (student level, adjusted for school-level clustering)

Across these eight studies that investigated the relationships of professional development to students' achievement, few significant effects were identified. The study that yielded the most consistently positive effects was Chapin (1994). In this study, the achievement of 723 third-, sixth-, and seventh-grade students of 42 teachers who participated in a professional development program was compared with students in the same schools whose teachers were not involved in the program. The study design does not permit analysis of particular features of the professional development that might account for differences in teacher performance. Two other studies that showed positive effects on student achievement were Carpenter, Fennema, Peterson, Chiang, and Loaf (1989) and Saxe, Gearhart, and Nasir (2001). Carpenter et al. studied a professional development program in which teachers were provided with knowledge of students' number fact concepts and reasoning, and showed that students in experimental classes exceeded students in control classes in number fact knowledge and problem solving. Saxe et al. compared three different professional development programs: an extensive subject-matter-focused professional development program whose goal was to improve teacher understanding of content, and of student thinking and learning processes; a program focused on building collegial support; and a traditional in-service program.[20]

The researchers found significant differences for pupils' conceptual and computational outcomes with the subject-matter-focused program leading to greater effects on students' conceptual learning and the traditional program yielding greater impact on students' computational skills. Angrist and Lavy (2001) studied a professional development program that involved weekly meetings between trainers and teachers to review teaching methods and plans for the following week, based on a "humanistic mathematics" philosophy of teaching. Treatment schools received an average of 10.5 more hours per week in training than comparison schools. Students in the treatment classes where the intervention was

[20] "Subject-matter-focused" denotes a program in which mathematics is central to teachers learning opportunities, and "traditional" ones are more focused on pedagogical classroom processes.

implemented performed significantly better than students in control schools. Still, none of these studies offers clear signals about the features of professional development that affect teachers' capacity to teach or students' achievement gains.

Only one study (Jacob & Lefgren, 2002) directly investigated effects of the amount of professional development, a factor frequently thought to be crucial for effective professional development. This study reported no effects from modestly increased amounts of professional development on students' learning in schools that were placed on probation for poor achievement. Little can be concluded from this study as the professional development was not specifically focused on mathematics, and no detailed information was provided about the nature of this professional development other than that it varied widely and was administered by a variety of organizations.

Overall, the Task Group was not able to draw conclusions about the features of professional development that have an impact on students' achievement because of the paucity of studies that investigated this link. For the studies the Task Group did identify that probed this connection, specificity is lacking regarding the features of the professional development programs where effects were found.

Professional development is often regarded as one of the key policy levers for improving instruction and student achievement with currently practicing teachers. To probe this assumption, the Task Group sought to identify and review the available evidence about effects of professional development, and features that make a difference for either student outcomes or their teachers' capacity to teach. The Task Group uncovered no studies, however, of sufficient quality where the designs and measures permitted them to ask and answer questions about teachers' learning. Most studies used a simple pre- and posttest design with no comparison group or used self-report data on teachers' learning. To ascertain the impact of professional development on students' achievement, the Task Group did identify a small number of studies, but overall, these did not support any specific claims about the nature of professional development that affects teachers' effectiveness.

5. Conclusions

The Task Group reviewed research on teacher education, including preservice teacher preparation, alternative pathways into teaching, induction programs, and professional development for practicing teachers. Despite the many beliefs about effective teacher education in any of these forms, the Task Group did not find strong evidence for the relationships between teacher education, and either teachers' capacity to teach or their students' learning. Even for the few studies that did produce significant effects, so little was unpacked about the features of the training that might account for a program's impact that the Task Group was left without much greater insight into the crucial components of teacher education.

Note that most research in this domain lacks rigorous designs and measures and is descriptive more often than not. Without comparison groups, or designs that permit analysis of program effects, it is difficult to draw conclusions about how teacher education works or about what the key features of effective professional training are.

6. Recommendations

Studies are needed that use designs that lead to knowledge about the impact of different approaches to professional development and permit comparisons with other potential impacts on teacher capacity and their effectiveness (e.g., experience, curriculum, curriculum policy). Such research will depend not only on rigorous designs but also on valid and reliable measures of the key outcome variables: teachers' mathematical knowledge and skill, instructional quality, and student learning. Self-reported data cannot continue to be the main source of information about professional development outcomes.

Key questions on which robust evidence is needed include the following:

- Does teacher education (e.g., preservice training of different kinds, professional development, early career induction programs) have an impact on teachers' capacity to teach and on students' achievement?
- What are key features of teacher education (e.g., duration, structure, quantity, content, pedagogy, structure, relationship to practice) that have effects on teachers' capacity to teach and on students' achievement?
- How do contexts (e.g., school, students, teachers, policy) affect the outcomes of professional development?
- How do different amounts of teacher education affect outcomes and effects?

Given the vast investment made in teacher education and the call for more of it, knowledge about its effects is vitally needed. Efforts to build measures and to implement better research designs should be supported.

C. Teacher Incentives

What types of recruitment and retention strategies are used to attract and retain highly effective teachers of mathematics? How well do they work?

As the Task Group has previously noted, substantial differences in the mathematics achievement of students are attributable to differences in teachers. The Task Group has focused on the role of teachers' mathematical knowledge in predicting student achievement, with teachers' knowledge measured by proxies, such as certification, college course work, and scores on tests. Teachers' college course taking in mathematics is associated with student gains in high school. Teacher's mathematical knowledge as measured by tests is also linked to student achievement. Other characteristics of teachers including years of teaching experience, general cognitive ability, and selectivity of the institution awarding the teacher's baccalaureate degree are associated with teacher effectiveness (Greenwald, Hedges, & Laine, 1996). It is important to note, however, that the largest reported positive effect of any of these characteristics is small relative to the magnitude of the natural variation in teacher effectiveness.

Thus the Task Group knows that there are large differences in the on-the-job performance of teachers of mathematics as measured by student gains. Some portion of the differences in teacher effectiveness can be predicted by such known characteristics of teachers as their college course work and their scores on tests. But prior on-the-job performance of teachers is by far the strongest predictor of their future on-the-job performance.

1. Utilizing Labor Market Incentives for Good Teaching

One possible mechanism for recognizing and leveraging differences in teaching ability is salary. According to the National Center for Education Statistics' Schools and Staffing Survey (Gruber, Wiley, Broughman, Strizek, & Burian-Fitzgerald, 2002), 70% of public school teachers in the United States in 2000 worked under a "uniform salary schedule" in which teachers with the same number of years of employment and the same level of postsecondary education received the same pay. In private industry and higher education, in contrast, pay is typically contingent on performance and area of specialization, as well as years of experience and level of education. In universities, for example, economists typically receive higher salaries than historians, reflecting the greater demand for economists outside the university sector. Within academia, economists who publish more influential work and bring in more external funding are paid more than less productive economists. Parallels in K–12 education would take the form of paying more to teachers who have technical skills that are in demand in other sectors of the economy, such as teachers with degrees in mathematics (*skills-based pay*), and paying more to mathematics teachers who are more productive in raising student achievement (*performance-based pay*). Another type of incentive intends to compensate teachers for working in conditions they view as unfavorable, such as those associated with high-poverty, low-achieving schools (*location pay*).

a. Skills-Based Pay

Skills-based teacher incentives are based on two premises: Certain types of preparation and training are necessary to teach certain subjects, and individuals with that preparation and training are less likely to enter into or remain in teaching if their levels of compensation are substantially below market rates. Evidence presented previously is consistent with the first premise in demonstrating a relationship between certain types of teacher preparation and student outcomes. With respect to the second premise, a large and consistent body of economic research indicates that college students' decisions to prepare for and enter into teaching depend on how the salary structure for teachers compares with those in competing occupations (Dolton & van der Kaauw, 1995; Bacolod, 2007; Goldhaber, DeArmond, Liu, & Player, 2007).

The magnitude of the salary differential between the private sector and the teaching profession for those who enter teaching with technical training is large. Four years after graduation, the gap in annual salary between teachers and non-teachers who have training in math and science is $13,469. Ten years out of college, the annual salary gap is $27,890 (Goldhaber et al., 2007). The salary differential for mathematically trained teachers versus non-teachers is likely to account at least in part for the significant teacher shortage in mathematics. In 2003–04, 74.1% of public high schools reported having teaching vacancies in mathematics, and 32.4% of those schools indicated that it was very difficult, or they were not able to fill those vacancies. Both of these levels were higher than for any other field in which high schools reported vacancies (Strizek, Pittsonberger, Riordan, Lyter, & Orlofsky, 2006). Differential salaries may also be responsible in part for the higher attrition rate from teaching of teachers with training in mathematics and science; Ingersoll (2000) reports that math and science teachers are significantly more likely to move from or leave their teaching jobs because of job dissatisfaction than are other teachers (40% of math and science, and 29% of all teachers). Of those who depart because of job dissatisfaction, the most common reason given is low salaries (57% of respondents).

b. Location Pay

"Location" pay is premised on the well-documented tendency of the most qualified teachers to select or migrate towards schools that serve the most economically advantaged children (Lankford, Loeb, & Wyckoff, 2002; Hanushek, Kain, & Rivkin 2004; Reed, Rueben, & Barbour 2006). Could this problem be remedied by paying teachers more who serve in high-needs schools? Research on the effects of location pay provides mixed results that are likely affected by the size of differential pay, the gender and experience of the teacher, and whether the bonus is a one-time signing bonus or permanent, among other factors (Hanushek et al.; Loeb & Page, 2000). Hanushek et al., using longitudinal data from Texas and taking advantage of naturally occurring variation in teacher salaries across districts, found that women were much less responsive to salary differences than men in determining whether to transition out of a high-minority school. They estimate that it would require an 8.8% salary premium for nonminority males with 3–5 years of teaching experience to keep them from moving from large urban to suburban districts, but a 42.6% differential to retain nonminority females. On the assumption that location pay could not be targeted to male teachers, they concluded that offering teachers pay differentials to take jobs in low-performing schools is not a cost-effective means of improving achievement. In contrast, Clotfelter, Glennie, Ladd, and Vigdor (2006) found that a moderately-sized addition to salary ($1,800) was effective in encouraging mid-career and more senior math and science teachers to stay in high-needs districts in North Carolina.

One important difference in the two studies is that Hanushek et al. (2004) estimated the size of the incentive that would neutralize teacher movement out of high-poverty urban schools, whereas Clotfelter et al. (2006) estimated the effect of the particular $1,800 bonus used in North Carolina, which was a 12% reduction in turnover rates. It may require much lower levels of location pay to reduce the outflow of experienced teachers from high-needs schools than it would take to eliminate it.

c. Performance Pay

Both skills-based pay and location pay as currently conceptualized and implemented provide incentives based on characteristics of teachers, such as college course work and experience, which are relatively weak predictors of student achievement. Thus they may be relatively inefficient mechanisms for enhancing the supply of effective teachers, where "effective" is defined as a teacher's above-average ability to increase the measured academic achievement of students. If, as previously documented, the strongest predictor of a teacher's effectiveness is the teacher's history of effectiveness, and if teachers' performance is affected by salary, then pay-for-performance might generate greater yields than similar investments in other incentives.

What is the evidence that pay-for-performance, or merit pay, has positive effects on teaching quality in mathematics? Before addressing that question it is important to note and briefly describe the substantial variability in the design of extant merit-pay systems.

One major categorical distinction is between incentive schemes focused on individuals versus those focused on schools. In the latter, teachers within schools receive a bonus if the entire school in which they teach makes progress on measures of student learning. In the former, individual teachers receive salary increments based on the gains of their own students over the course of the school year. School-based incentive plans typically have political advantages over individual plans, and some have argued that they enhance cooperation among teachers compared to individual plans. But others have argued that school-based plans, depending as they do on cooperation and continuity of effort by the teaching staff, are poorly designed for schools in which teacher mobility is high, a striking characteristic of many central city, high-poverty schools.

A second important dimension of variation among pay-for-performance plans is the level of compensation bonus that is available for higher performing teachers. Existing research on pay-for-performance involves bonuses that range from a couple of hundred dollars (Lavy, 2002) up to 40% or more of base salary (Glewwe, Ilias, & Kremer, 2003).

Another distinction involves the degree to which the merit system is focused on student outcomes. In some systems, teacher bonuses are totally dependent on student gains on standardized assessments of learning (e.g., Winters, Ritter, Barnett, & Greene, 2006). At the other end of the dimension are systems in which input and output measures are mixed together in complex arrays. An example is Mexico’s Carrera Magisterial, which rewards teachers with salary bonuses based on a number of criteria, such as seniority, educational attainment, professional development, teacher performance, and student achievement (Santibáñez et al., 2007).

Continuity is yet another dimension on which pay-for-performance systems differ. The majority of existing pay-for-performance systems is put forward as trials of the concept. A system that seems temporary may motivate teachers differently than one around which longer-term plans can be made.

The Task Group identified 14 quantitative studies on teacher merit pay (see Table 13). Of these, 13 showed positive effects on student outcomes (see Table 14). These studies varied methodologically from randomized control trials (Muralidaran, & Sundararaman, 2006; Glewwe et al., 2003) to causally weaker correlational studies (e.g., Figlio & Kenny, 2007) and quasi-experiments (e.g., Ladd, 1999). The design of the performance pay plans varied across the studies on all the design features identified in this section. Given the variability in program design and evaluation methodology, it is striking that each of the studies found some positive effects on student achievement.

Table 13: Quality Characteristics of Studies Examining Impact of Teacher Pay for Performance on Student Achievement in Mathematics

			Other Controls							
Authors	**Sample Size**	**Pretest Control**	**Student**	**Teacher**	**Class, School, or District**	**Family**	**Matching of Schools**	**Level of Analysis**	**Incentive Scheme**	**Math Specific Test**
Atkinson et al., 2004	182 secondary school teachers in the U.K. across 18 schools (with 23,000 students)	X	X	X				Teacher	Teacher-based: greater than 9% increase in salary	X
Cooper, & Cohn, 1997	541 classes, comprised of 532 teachers and 13,646 students in 18 school districts	X		X	X	X		Class/Teacher	2 plans—teacher-based or teacher- and school-based: each with similar pay incentives from $2,000–$3,000	X
Dee, & Keys, 2004	24,000 observations pooled from over 11, 000 students in 79 participating schools over 4 years	X	X	X	X	X		Student	Teacher-based: $2,000–$7,000 raise	X
Eberts, Hollenbeck, & Stone, 2002	2 high schools in MI (collectively comprised of over 17,000 student/year observations)	X					X	School	Teacher-based: student retention bonus 12–12.5% of base pay, performance bonus 5% of base pay, additional 10% bonus for both	
Figlio, & Kenny, 2006	4,515 students	X	X		X	X		Student	Varied within sample	
Glewwe et al., 2003	Over 50,000 students from 100 primary schools	X	X		X			Student	School-based: 21–43% of base salary	X
Ladd, 1999[a]	1,118 school-year observations in 5 cities				X			School	School-based: $1,000 to teacher and principals, and $2,000 to schools	X
Lavy, 2002	Over 22,000 observations across 190 schools		X		X	X	X	Student	2 school-based plans: amount varies by plan	X
Lavy, 2004	Over 120,000 student-year observations across 350 schools		X		X	X	X	Student	Teacher-based: $1,750–$7,500+	X
Muralidharan, & Sundararaman, 2006	Over 68,000 students from 300 schools	X			X	X		Student	School-based, teacher-based: Unrestricted. 500% gain in ave. student test score – 5%)	X

Continued on p. 5-46

Table 13, continued

Authors	Sample Size	Pretest Control	Other Controls: Student	Teacher	Class, School, or District	Family	Matching of Schools	Level of Analysis	Incentive Scheme	Math Specific Test
Richards, & Sheu, 1992	994 schools	X						School	School-based: amount varies, during study was \$30–\$50 per pupil or approximately \$10,000 per school	
Santibáñez et al., 2007	Over 850,000 classroom-year observations in Mexico			X	X	X		Class/Teacher	(MX\$)[b] 1,599 to (MX\$) 12,462. Representing 27–215% of base salary	
Slotnik, Smith, Glass, & Helms, 2004	Over 100,000 student-year observations across 16 schools in CO		X		X	X	X	Student	\$500–\$1,500	X
Winters, Ritter, Barnett, & Greene, 2006	608 student-year observations across 5 schools in AR	X	X		X	X	X	Student	\$1,800–\$11,200	X

[a] Another piece, a book chapter by Clotfelter and Ladd (1996), has been cited by such authors as Podgursky and Springer (2007). However this piece has been omitted from Tables 13 and 14 because the analysis in that chapter later became integrated as part of a published article by Ladd (1999), and presents the same analysis as the Ladd (1999) piece included here.
[b] MX\$ means Mexican dollars.

The methodologically strongest studies have been conducted in developing countries. For example, Muralidaran and Sundararaman (2006) reported results from a randomized trial of individual- and school-level performance-based incentives implemented across a representative sample of government-run rural primary schools in the Indian state of Andhra Pradesh. The program provided bonus payments to teachers based on the average improvement of their students' test scores in independently administered learning assessments (with a mean bonus of 3% of annual pay). The effect size for students in incentive schools was .19 standard deviation units for mathematics. The students scored significantly higher on "conceptual" as well as "mechanical" components of the test suggesting that the gains in test scores represented an actual increase in learning outcomes. Incentive schools also performed better on subjects for which there were no incentives. There was no significant difference in the effectiveness of group versus individual teacher incentives. Incentive schools performed significantly better than other randomly chosen schools that received additional paraprofessional teachers and cash block grants that were equivalent in costs to the teacher incentives.

No experiments on performance-based pay have been reported in the United States, although one large randomized trial is underway in Nashville, TN by the National Center on Performance Incentives. A recent correlational study by Figlio and Kenny (2006) examined locally generated, individual-based merit-pay programs in the United States by combining data from the NELS with the authors' own survey on the use of incentives. The performance plans varied from school to school. The authors found that merit-pay plans had positive impacts on student achievement and appeared to be effective when other types of interventions, such as more frequent teacher evaluation, were not.

2. Recommendations

The results from research on teacher incentives generally support the effectiveness of incentives, although the methodological quality of the studies in terms of causal conclusions is mixed. The substantial body of economic research in other fields indicating that salary affects the number of workers entering a field and their job performance is relevant. In the context of the totality of the evidence, and acknowledging the substantial number of unknowns, the Task Group recommends policy initiatives that put in place and carefully evaluate the effects of the following:

- Raising base salaries for teachers of mathematics to be more competitive with salaries for similarly trained non-teachers;
- Incentives for teachers of mathematics working in locations that are difficult to staff; and
- Opportunities for teachers of mathematics to increase their base salaries substantially by demonstrable effectiveness in raising student achievement.

3. Cautions

The lack of results from randomized trials of performance-pay systems in the United States and the difficulty of estimating a cost-benefit ratio for particular types of bonuses means that much work remains to be done before the nation will know enough to put particular pay-for-performance systems in place and predict their outcomes confidently. Currently, the effects of pay-based incentives on teachers are largely unknown. Knowing more about how various incentive systems affect teachers would enable the design of more effective and efficient incentives. Beyond the uncertainties about the effects of particular incentive systems, there is substantial evidence that teachers' decisions to remain in teaching and to continue teaching in particular schools are affected by work conditions in addition to salary. This includes the proximity of their residence to the school, their support from school administrators, their teaching assignment, and the characteristics of their students (Marvel, Lyter, Peltola, Strizek, & Morton, 2006; Hanushek et al., 2004). It is important to note that increasing the pay of mathematics teachers necessarily involves redirecting resources from other purposes, including investments that might have greater effects on student outcomes. Informed policy decisions need to take into consideration the relative returns of alternative investments in improving student achievement. In light of the substantial number of unknowns, policy initiatives involving teacher pay should be carefully evaluated as they are put in place.

Table 14: Reported Impacts of Models Examining the Effect of Teacher Pay for Performance on Student Achievement in Mathematics

Authors	Dependent Measure	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade(s)	Estimation Technique
Atkinson et al., 2004	Gains in GSCE mean math test score	Teacher's eligible for incentives	Data not available	Positive, Significance unknown	X	8–10	OLS regression
	Gains in (KS3 - GSCE) math test scores	Teacher's eligible for incentives	Data not available	Positive, Significance unknown	X	8–10	OLS regression
Cooper, Cohn, 1997	Gains in median math test score	Participation in Teacher Bonus Model	0.22	Positive & Significant	X	HS & Elem.	OLS regression
	Gains in median math test score	Participation in Campus/ Individual Incentive Model	0.13	Positive & Significant	X	HS & Elem.	OLS regression
	Gains in median math test score	Participation in Teacher Bonus Model	0.20	Positive & Significant	X	HS & Elem.	Frontier regression
	Gains in median math test score	Participation in Campus/ Individual Incentive Model	0.13	Positive & Significant	X	HS & Elem.	Frontier regression
Dee, & Keys, 2004	Stanford Achievement Test score in math	Participation in TN Career Ladder System	Data not available	Positive & Significant	X	K–3	OLS regression
Eberts, Hollenbeck, & Stone, 2002	Course completion	Implementation of performance pay incentives in MI	Data not available	Positive & Significant		HS	Difference-in difference in mean outcomes
	Student GPA	Implementation of performance pay incentives in MI	Data not available	Negative & Not significant		HS	Difference-in difference in mean outcomes
	Course-passing rates conditional on course completion	Implementation of performance pay incentives in MI	Data not available	Negative & Significant		HS	Difference-in difference in mean outcomes
Figlio, & Kenny, 2006	Sum of 12th-grade NELS:88 scores across subjects	Number of Teacher incentives offered	Data not available	Positive & Significant		12	OLS regression
	Sum of 12th-grade NELS:88 scores across subjects	The existence of large incentives for teachers	Data not available	Positive & Significant		12	OLS regression
	Sum of 12th-grade NELS:88 scores across subjects	The existence of medium incentives for teachers	Data not available	Positive & Significant		12	OLS regression
	Sum of 12th-grade NELS:88 scores across subjects	The existence of small incentives for teachers	Data not available	Positive & Significant		12	OLS regression

Continued on p. 5-49

Table 14, continued

Authors	Dependent Measure	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade(s)	Estimation Technique
Glewwe et al., 2003	Change in district exam scores in math	First year as incentive school	0.04	Positive & Not significant	X	4–8	GLS regression in a random effects model
	Change in district exam scores in math	Second year as incentive school	0.08	Positive & Significant	X	4–8	GLS regression in a random effects model
	Change in district exam scores in math	First year after two-year incentive program	-0.05	Negative & Not significant	X	4–8	GLS regression in a random effects model
	Change in KCPE exam scores in math	First year as incentive school	0.06	Positive & Not significant	X	4–8	GLS regression in a random effects model
	Change in KCPE exam scores in math	Second year as incentive school	0.07	Positive & Not significant	X	4–8	GLS regression in a random effects model
	Change in KCPE exam scores in math	First year after two-year incentive program	0.04	Positive & Not significant	X	4–8	GLS regression in a random effects model
Ladd, 1999[a]	Pass rates in math on the TX Assessment of Academic Skills (TAAS)	First year of incentive reform	Data not available	Positive & Significant	X	7	OLS regression
	Pass rates in math on the TX Assessment of Academic Skills (TAAS)	Second year of incentive reform	Data not available	Positive & Significant	X	7	OLS regression
	Pass rates in math on the TX Assessment of Academic Skills (TAAS)	Third year of incentive reform	Data not available	Positive & Significant	X	7	OLS regression
	Pass rates in math on the TX Assessment of Academic Skills (TAAS)	Fourth year of incentive reform	Data not available	Positive & Significant	X	7	OLS regression
Lavy, 2002	Ave. matriculation test scores in math, in secular schools	Two years after adopting incentive program	Data not available	Positive & Significant	X	12	OLS regression using regression discontinuity
	Ave. matriculation test scores in math, in religious schools	Two years after adopting incentive program	Data not available	Positive & Significant	X	12	OLS regression using regression discontinuity
	Ave. matriculation test scores in math, in secular schools	Two years after adopting incentive program	Data not available	Positive & Significant	X	12	OLS regression using RDD & PS matching
	Ave. matriculation test scores in math, in religious schools	Two years after adopting incentive program	Data not available	Positive & Significant	X	12	OLS regression using RDD & PS matching

Continued on p. 5-50

Table 14, continued

Authors	Dependent Measure	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade	Estimation Technique
Lavy, 2004	Awarded credits in math	Adoption of incentive program	Data not available	Positive & Significant	X	12	Regular OLS regression
	Passing rates in math on matriculation exams	Adoption of incentive program	Data not available	Positive & Significant	X	12	Regular OLS regression
	Awarded credits in math	Adoption of incentive program	Data not available	Positive & Significant	X	12	OLS regression using PS matching
	Passing rates in math on matriculation exams	Adoption of incentive program	Data not available	Positive & Significant	X	12	OLS regression using PS matching
	Awarded credits in math	Adoption of incentive program	Data not available	Positive & Significant	X	12	OLS regression using RDD
Muralidharan, & Sundararama, 2006	Difference in avg. math test designed by “educational initiatives”[b]	School assignment to individual incentives	Data not available	Positive & Significant	X	1–5	OLS regression
	Difference in avg. math test designed by “educational initiatives”	School assignment to group incentives	Data not available	Positive & Significant	X	1–5	OLS regression
Richards, & Sheu, 1992	School mean gain score on Basic Skills Assessment Program (BSAP)	Rank or “band” within established SC incentive structure	Data not available	Positive, Significance unknown		1–11	Comparison of means
	School mean gain score on Comprehensive Test of Basic Skills (CTBS)	Rank or “band” within established SC incentive structure	Data not available	Positive, Significance unknown		1–11	Comparison of means
Santibáñez et al., 2007	Carrera Magisterial test score (across many subjects)	IP Score, i.e. incentive scheme ranking metric	0.16	Positive & Significant		Primary	OLS regression using RDD
	Carrera Magisterial test score (across many subjects)	Closest to cutoff for receiving incentives	0.00	Negative & Not significant		Primary	OLS regression using RDD
	Carrera Magisterial test score (across many subjects)	IP Score, i.e. incentive scheme ranking metric	0.07	Positive & Significant		Secondary	OLS regression using RDD
	Carrera Magisterial test score (across many subjects)	Closest to cutoff for receiving incentives	0.01	Positive & Not significant		Secondary	OLS regression using RDD

Continued on p. 5-51

Table 14, continued

Authors	Dependent Measure	Independent Variable	Standardized Regression Coefficients	Results of Estimated Effects	Math Specific Outcome	Grade	Estimation Technique
Slotnik, Smith, Glass, & Helms, 2004	Iowa Test of Basic Skills (ITBS)	Marginal effect of being a pilot school (PFP plan adopted)	Data not available	Negative & Significant	X	Elem.	2 stage HLM model
	Colorado Student Assessment Program (CSAP)	Marginal effect of being a pilot school (PFP plan adopted)	Data not available	Negative & Significant	X	Elem.	2 stage HLM model
	Iowa Test of Basic Skills (ITBS)	Marginal effect of being a pilot school (PFP plan adopted)	Data not available	Positive & Not significant	X	Middle	2 stage HLM model
	Colorado Student Assessment Program (CSAP)	Marginal effect of being a pilot school (PFP plan adopted)	Data not available	Positive & Significant	X	Middle	2 stage HLM model
	Iowa Test of Basic Skills (ITBS)	Marginal effect of being a pilot school (PFP plan adopted)	Data not available	Positive & Significant	X	HS	2 stage HLM model
	Colorado Student Assessment Program (CSAP).	Marginal effect of being a pilot school (PFP plan adopted)	Data not available	Negative & Not significant	X	HS	2 stage HLM model
Winters, Ritter, Barnett, & Greene, 2006	NCE in math from SAT minus NCE in math from ITBS	Participation in incentive program	Data not available	Positive & Significant	X	K–5	OLS regression, Difference-in-difference

[a] Another piece, a book chapter by Clotfelter and Ladd (1996), has been cited by such authors as Podgursky and Springer (2007). However this piece has been omitted from Tables 13 and 14 because the analysis in that chapter later became integrated as part of a published article by Ladd (1999), and presents the same analysis as the Ladd (1999) piece included here.

[b] Educational Initiatives refers to incentives given or provided to educators based on achievement or attainment.

D. Elementary Mathematics Specialist Teachers

a) *What models exist for elementary math specialist teachers and their preparation? What is known about the qualifications and responsibilities of mathematics specialist teachers?*
b) *What evidence exists for the effectiveness of elementary math specialist teachers with respect to student achievement?*

There have been many calls for the use of math specialists at the Grade K–5 level in recent years (National Research Council, 2001; Maryland State Department of Education, 2001; National Council of Teachers of Mathematics, 2000; Horowitz, Fuchs, & Clarke, 2006; American Mathematical Society, 2001; Fennell, 2006). The pressing need of math specialists comes from two sources. On the one hand, there is now a general awareness that many elementary teachers lack adequate knowledge of mathematics for teaching. Moreover, evidence exists for substantial variability in teachers' knowledge of mathematics for teaching, and evidence exists that teachers' grasp of such knowledge is directly and very strongly related to the mathematical quality of their classroom instruction (Learning Mathematics for Teaching, 2006). On the other hand, with more than 2 million current elementary teachers, the scale problem of raising the mathematical knowledge of such a large number of teachers becomes intractable. The hope of training a small cadre of mathematically knowledgeable teachers and letting them teach the elementary mathematics classes leads some to consider the use of math specialists an important and scalable route to improving the quality of the mathematics instruction that students receive. The use of math specialists at the Grade K–5 level to reduce the number of teachers who must know mathematics well for teaching therefore seems like a sensible strategy.

However, despite multiple recommendations to use math specialist teachers, the meaning of the term "math specialist" varies. While educators and administrators who carry the title of "math specialist" can be found in most states, they frequently have different sets of responsibilities and qualifications. No national survey of their numbers, responsibilities, or qualifications has been conducted. Recent surveys of math specialists in Iowa[21] and Maryland[22] have shown that the use of math specialists is widespread, and in the case of Maryland this is true across multiple models of the position. Not long ago the Virginia legislature mandated the use of math specialists across the state; the state currently has one of the most developed and researched math specialist programs. Similar legislation may soon be proposed in the state of Maryland (Wray, 2007). Washington has also adopted legislation funding the use of math specialists to act as teacher coaches across the state (Thompson, 2007). The use of different models of math specialists is thus by no means uncommon, but as the following discussion shows, their presence in the education hierarchy has been poorly defined and their effectiveness has been insufficiently studied.

To contribute to thoughtful consideration of the issues involved in restructuring teacher roles around the idea of mathematics specialists, the Task Group reviewed a range of models in current use in the United States and abroad, and sought evidence about their effectiveness. However, there is a paucity of rigorous empirical research to answer the question posed in this section.

1. What Models Exist for Elementary Math Specialist Teachers and Their Preparation?

Math specialists can be found working at every level of U.S. public school systems. They hold positions that oversee all or groups of districts within a state, a single district, a single school within a district, classrooms within a school, and even particular students within a classroom. Some math specialists even take on several of these duties all at once (W. Haver, personal communication, April 1, 2007). In middle schools, math specialists are

[21] A survey of superintendents in Iowa shows that 63.2% of elementary schools in the state have some type of departmental model that treats math as a separate subject requiring a math specialist working under what likely is the specialized-teacher model. Not all elementary grades worked under this model. However, only 2.7% of Grades 1 and 2 were departmentalized across all elementary schools. This compares to 9% of all Grades 3, 46.8% of Grades 4, 89.2% of Grades 5, and 66.7% of Grades 6 that were departmentalized across all elementary schools (Kemis, Heiting, Spitzli, & Lang, 2003).

[22] "In a recent survey of Maryland's local mathematics supervisors, data indicated that Maryland public schools (Grades PreK–12) currently employ approximately 445 school-based (non-teaching) mathematics teacher-leadership specialists at 439 schools, 158 school-based mathematics intervention specialists (who work primarily with students), and 134 district-level mathematics curriculum/instruction specialists, statewide" (Ruehl & Wray, 2006).

employed most often specifically to teach mathematics, while in elementary schools they may teach their students about multiple subjects, not just mathematics (Fennell, 2006). Overall, the location of employment of math specialists within the structure of a state's school system partially determines both the certification needed to qualify for the position and the responsibilities of the position.

Different sets of qualifications are currently required of mathematics specialists. Minimally, a math specialist is someone who has demonstrated expertise, usually through experience as a former math teacher, testing, or both.[23] Math specialists are frequently certified through course work leading to a master's degree (W. Haver, personal communication, April, 2007). For example, particular colleges and universities in the states of Virginia[24] and Michigan[25] have graduate-level certification programs for mathematics specialists. Math specialists have also been known to gain their certification through professional development course work not leading to a higher degree from any accredited institution of higher or postsecondary education (C. Chapman, personal communication, May 1, 2007).

At present, there are at least three models for the use of math specialists: the lead-teacher or math coach model, the specialized-teacher model, and the pull-out model (Fennell, 2006). Math specialists as lead teachers are more common than those working under the other two models. In practice, however, math specialists frequently take on responsibilities that cut across all three models.

[23] For example, the governor of the state of Massachusetts recently recommended laws requiring math specialists to be elementary school teachers who have spent at least 80% of their weekly teaching time teaching math and who also have passed the elementary math section of the Massachusetts Tests for Educator Licensure (Romney, 2007).

[24] There are currently six colleges and universities in the state that offer a master's degree program that endorses mathematics specialist in accordance with states' licensure regulations for school personnel (Virginia Legislature, 2005).

[25] "In conjunction with the Horace H. Rackham School of Graduate Studies, the School of Education of the University of Michigan-Dearborn offers a Master of Arts degree in Education. This is a degree that is designed for educators who desire to fulfill all requirements for a University of Michigan master's degree, including residency, at UM-D. The program is also designed for teachers who wish to strengthen their competencies, expand their professional outlook and gain greater knowledge and understanding of their subject specialization. Through this program teachers may apply for an endorsement/certificate in Early Childhood and Early Childhood Special Education Inclusion UM-D Certificate, English as a Second Language, Middle Level Education, Middle School Mathematics Endorsement and Middle Grades Mathematics Leadership, Reading Specialist K–12, other endorsements for which the School of Education is approved, a renewal of the Provisional Certificate, or obtain a Professional Education Certificate. Eligibility for regular admission into the program includes completion of a bachelor's degree, a 3.0 (B) undergraduate grade point average or better, and a teaching certificate. Individuals whose grade point is less than 3.0 may be considered for probationary admission status and may be required to submit evidence of potential for success in a graduate program." (Retrieved on 8/13/2007 from http://www.umich.edu/~bhlumrec/acad_unit/rackham/degree_req/ www.rackham.umich.edu/Programs/other.camp/Dearborn/educ-dbn.html.)

a. The Lead Teacher or Math Coach Model

Math specialists of this type act as resource persons for their coworkers and do not directly instruct students. They work at the state, district, and school levels, providing leadership and information to teachers and staff and often coordinating mathematics programs within a school, a district, or across districts (V. Inge, personal communication, May 1, 2007). Frequently, these math specialists are elementary teachers who have been released from all or some of their responsibilities of classroom instruction (Reys & Fennell, 2003; Rowan & Campbell, 1995; Ohanian, 1998). They facilitate teachers' use of instructional strategies; align curricular frameworks with state, district, or local standards; distribute, interpret, and apply the findings of research on the teaching and learning of math to teachers; and organize professional development opportunities for teachers (M. Madden, personal communication, April 1, 2007). One of the most important responsibilities of math specialists of this type is their charge to foster a self-reflective culture of learning among teachers (Moon, 2002; P. Hess, personal communication, May 1, 2007). As Campbell (2007; Campbell & White, 1997) points out, theirs is a very challenging task as,

> [math specialists of this type] are called upon to navigate not only the complexity of teaching and student learning as it emerges in the classrooms of multiple teachers, but to do so while provoking the development of those teachers by advocating for their change, nurturing their performance, advancing their thinking, increasing their mathematical understanding, and saluting their attempts (Campbell & White, 1997, p. 328).

Some math specialists are also empowered to officially assess the performance of math teachers in the classroom, but this role is less common as it requires particular organizational structures within schools or districts to support this form of professional evaluation.[26] Math specialists of this type have been funded in part with Title I dollars.[27] In practice, the responsibilities of math specialists who work as lead teachers are quite fluid and are primarily determined by the context of their employment. They are referred to by many other names, including mentors, coaches, and resource, lead, and peer teachers (J. Lott, personal communication, May 1, 2007). As with their responsibilities, the meaning of these terms also depends on context (Miller, Moon, & Elko, 2000).

---

[26] For instance, the Charlottesville City Schools District in Charlottesville, VA, has hired elementary math specialists to work in individual schools who supports "the professional growth of elementary teachers by strengthening classroom teachers' understanding of math content … will co-teach lessons, develop curriculum and lessons, and create appropriate assessments, as needed. The math specialist will spend a minimum of 75% to 80% of his/her time in classrooms directly with students. He or she may also assist administrative and instructional staff in interpreting data and designing approaches to improve student achievement and instruction. Qualified applicants must hold a valid Virginia teaching certificate and be enrolled in a university-based elementary math specialist program which will result in a Virginia endorsement as a K–8 math specialist. Applicants must have teaching experience and a master's degree is preferred" (Job posting found April 2007, at: http://www.ccs.k12.va.us/uploads/Mathematics%20Specialist%20JD.pdf).

[27] For instance, the town of Carlisle, MA, has used a Title 1 grant to partially fund a math specialist position in their elementary schools since 2003 (The Carlisle School Administration, 2005).

b. The Specialized-Teacher Model

Math specialists of this type are responsible for the direct instruction of students. They work at the school and district levels, but most frequently take responsibilities in one school. In the upper grade levels (particularly for Grades 6–8) these math specialists frequently have the responsibility for the instruction of a single grade level (Fennell, 2006). They work with other teachers at their grade level to divide up the subjects being taught (e.g., math, social studies, science), and frequently all teachers retain some responsibility for reading and language arts instruction in a "homeroom" environment (Reys & Fennell, 2003).

c. The Pull-Out Model

This is a variation of the specialized-teacher model. In this model, math specialists directly instruct individuals or small groups of students within a classroom who have been identified as either failing to meet or exceeding the standards attached to their grade level (V. Mills, personal communication, May 1, 2007). Therefore, this type of math specialists does not address the problem of the deficiency of mathematics instruction in the generic elementary classroom. They have been funded in part with Title I dollars, and in some cases Title II class reduction dollars (V. Inge, personal communication, May 1, 2007).

2. What Evidence Exists for the Effectiveness of Elementary Mathematics Specialist Teachers With Respect to Student Achievement?

The search for empirical investigations about the effectiveness of elementary mathematics specialists turns up little research on the subject. In all, the searches identified 114 potentially relevant pieces of literature, but only one (McGrath & Rust, 2002) was an investigation that explored the effects of specialized mathematics teachers on student achievement in elementary schools. These authors found no difference in the mathematics gain scores of students in an elementary school with a departmentalized structure compared to students in a school with a self-contained structure.

In the series of ethnographic studies compiled as part of the Recognizing and Recording Reform in Mathematics Education Project (R^3M), the presence of mathematics specialists or leaders were cited as being critical elements of reform (Ferrini-Mundy & Johnson, 1996). But the authors caution that the presence of the math specialists in these studies is not sufficient to make the argument for mathematics specialists.

Anecdotal reports addressing the effectiveness of math specialists are more common. These include descriptions of the use of math specialists working under the lead-teacher model (Campbell, 1996; West & Straub, 2003; Miller, Moon, & Elko, 2000) and the specialized-teacher model (Sconiers, 1991). Additional work presents anecdotal evidence showing that math specialists working under the lead-teacher model have a positive effect on students' academic achievement, or on their teachers' beliefs about teaching and learning, or on both (Brosnan, 2007; Campbell, 2007; Inge, 2002; Ruehl & Wray, 2006; Virginia Mathematics and Science Coalition, 2002; Virginia Legislature, 2006; Wray, 2007). In addition, Wu (2007) argues from a curricular perspective that the use of math specialists in the specialized-teacher model is a necessity for adequate mathematics instruction in elementary school.

Given the paucity of evidence that general teacher certification has a positive effect on student achievement, it may seem counterintuitive to think that the use of elementary mathematics specialists would have positive effects. It is likely, however, that if the use of elementary math specialists is to have a positive effect, it will be because the training of specialists develops, in a more focused way, the specialized mathematical knowledge for teaching shown to have effects on student achievement. This suggests that policies and programs for elementary math specialist need to be developed in tandem with research that attempts to uncover those aspects of teacher knowledge and understanding most strongly related to student learning.

a. Costs Associated With Mathematics Specialists

The costs associated with the employment, training, or certification of math specialists can be substantial.

Funding of employment. One widely held belief is that any commitment made to employ a math specialist must be maintained over time, usually for at least five years (Reys & Fennell, 2003). The costs quickly escalate when officials try to do so in multiple locations across a district or state.[28] These financial costs can change dramatically depending upon the type of math specialists employed. The lead-teacher model often requires a substantial commitment in resources because teachers may have to be reassigned to take on leadership roles. On the other hand, the specialized-teacher model may not require the hiring of new personnel but rather a redistribution of teaching tasks among existing employees and therefore could be less costly than the lead-teacher model (Reys & Fennell, 2003). The pull-out model may be more expensive than the specialized-teacher model because it requires additional teachers who target their instruction only at specific and small groups of students. Additionally, the use of the pull-out model may not rule out the need for employing multiple math specialists within one school.

Required training. Math specialists who work under the lead-teacher model may need instruction in math content, teaching pedagogy, education policy, organizational leadership, adult education, and professional development practices in education. While math specialists who work under the specialized-teacher model may not need training in organizational leadership, or adult education, or both, they do require the special training in mathematics necessary to directly instruct students with special needs (W. Haver, personal communication, April 1, 2007; V. Mills, personal communication, May 1, 2007).

[28] For example, the town of Carlisle, MA, hired one part-time math specialist at the cost of $25,000 a year in 2005 (The Carlisle School Administration, 2005). If math specialists are placed in every school in any given district, the cost quickly escalates. For example, in Delaware, a budget of $2.7 million for the math specialist program would provide a specialist for every school containing seventh and eighth grades. The governor of Alabama proposed the Math, Science and Technology Program—which provides professional development, equipment, and at-school support by math and science specialists to improve math and science instruction—that would allow the program to expand to more than 600 schools and serve more than 300,000 students at a cost $33 million. While it is impossible to extrapolate any real cost estimates from these figures, they do give a sense of the scale.

Training programs vary greatly in quality and length, and although a greater emphasis on the study of mathematics is needed to train current elementary teachers to be math specialists, simply taking more college-level mathematics courses would not be sufficient. In most cases, college-level mathematics courses are generally not well designed for the benefit of Grades K–6 teachers (Battista, 1994; Reys & Fennell, 2003). Many training programs do not require the completion of a master's degree, but rely on professional development opportunities, often lasting several days or weeks at least and frequently including some college-level course work.[29] The most rigorous training programs for math specialists require the completion of a master's degree, such as that found in the state of Virginia.[30]

Implementation. There are many costs associated with the adoption of any type of math specialist position. The scarce supply of qualified candidates available to work as math specialists plays a large role in driving up the costs associated with their use (Reys & Fennell, 2003). Additionally, other common barriers to the use of math specialists in elementary schools have been cited by researchers and practitioners.

Across most types of math specialists, there is currently little general consensus as to the appropriate certification required for employment and few common job descriptions. This can be considered a cost of implementation because an agreement on the necessary training and responsibilities of a math specialist may aid in the creation of necessary research and public policy (W. Haver, personal communication, April 1, 2007; L. Pitt. personal communication, May 1, 2007).

Differences between urban, suburban, and rural schools are so far-reaching and substantial that no single definition of the qualifications and responsibilities of math specialists can be easily adopted that would be valid across the board (L. Pitt, personal communication, May 1, 2007). For example, some rural areas use on-site teacher leaders as a means of offering leadership to small but spread-out populations of teachers and some urban districts are positioning several math specialists within one school (Campbell, 2007).

[29] The Alliance for Improvement of Mathematics Skills PreK–16 (AIMS PK-16) is a partnership of nine independent school districts in south Texas and two Hispanic-serving institutions of higher education, Del Mar (community) College and Texas A&M University-Kingsville. This partnership includes course work or organized collaboration with peers or experts in the subject of mathematics but does not lead to a degree from any accredited institution of higher or postsecondary education. (Retrieved on 8/13/2007 from http://www.delmar.edu/aims/)

[30] Virginia is home to six colleges and universities that offer a master's degree program for mathematics specialists. For example, the Mathematics Education Leadership Program at George Mason University offers a 33-credit master of education leadership degree with a concentration in math specialist leader (Grades K–8). The concentration is a unique 3-year program for persons who desire part-time study to become specialists in the teaching and leadership of school mathematics. Students in the program study mathematics content, teaching, curriculum, and leadership.

b. Mathematics Specialists Internationally

Eleven nations (Singapore, Belgium (Flemish), Sweden, Japan, China, the Netherlands, Latvia, Lithuania, United Kingdom, Hungary, and the Russian Federation) scored above the United States in the fourth-grade Trends in Mathematics and Science Study (TIMSS).[31] Only Singapore, Sweden, and China use math specialists in the specialized-teacher model in Grades 1 through 6 and the other seven do not; no data could be obtained from the Russian Federation. In summary, the utilization of math specialists in countries scoring higher than the United States on standardized tests, such as the TIMSS, is not widespread.

3. Conclusions

The lack of data precludes any definitive recommendations on the use of mathematics specialists. It may be noted, nevertheless, that the specialized-teacher model of math specialists is one that comes closest to the original intent of using math specialists: It is the least expensive of the three models, and it is the one among the three that seems the most realistic in solving the scale problem of overcoming the content-knowledge deficiency among elementary teachers. But the absence of data to support its potential effectiveness presents a problem in formulating policies for its widespread adoption. It might be a surprise that only 3 of the 11 nations that outperform the United States in the fourth-grade TIMSS use mathematics specialists. This is, however, difficult to interpret because in contexts where teachers' mathematical preparation is strong, the need for a subset of teachers selected to be specialists may be reduced. Still, there are compelling reasons to encourage research to examine the effectiveness of the specialized model. In addition, the need for this kind of math specialist may be sufficiently compelling so that one may wish to proceed, with caution, to create a corps of such specialists. In terms of content knowledge, the criteria that should be used for the certification of these specialists remain unknown.

The lead-teacher model is the most expensive of the three. It is also limited by the expectation that it is possible to produce a large number of teachers who not only possess superior mathematical knowledge to mentor teachers but also who are superior in pedagogical knowledge and organizational skills. Unless there is substantial evidence that this model of specialists is effective, any pursuit of this model may be premature at this point. To the extent that the pull-out model is not designed to meet the needs of the generic classroom, this model is not pertinent to the present considerations.

[31] Major sections of the research synthesis reported here were prepared by Institute for Defense Analyses Science Technology Policy Institute**.**

BIBLIOGRAPHY

Adams, P.E., & Krockover, G.H. (1997). Concerns and perceptions of beginning secondary science and mathematics teachers. *Science Teacher Education, 81*, 29–50.

American Mathematical Society (2001). *The mathematical education of teachers*. Washington, DC: Author.

Angrist, J.D., & Lavy, V. (2001). Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools**.** *Journal of Labor Economics, 19*(2), 343–369.

Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., & Wilson, D. (2004). *Evaluating the impact of performance-related pay for teachers in England.* Bristol, UK: Centre for Market and Public Organization, University of Bristol.

Bacolod, M. (2007). Who teaches and where they choose to teach: College graduates of the 1990s. *Educational Evaluation and Policy Analysis*. *29*(3), 155–168.

Battista, M. (1994). Teacher beliefs and the reform movement in mathematics education. *Phi Delta Kappan*, *75*(6), 462–470.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy, 1*(2) 176–216.

Bring, J. (1994). How to standardize regression coefficients. *The American Statistician*, 48 (3): 209–213.

Brosnan, P. (2007). *Mathematics coaching influence on teaching and learning middle school mathematics*. Unpublished presentation of preliminary findings, The Ohio State University, Mathematics Coaching Program, Columbus, OH.

Campbell, P.F. (1996). Empowering children and teachers in the elementary mathematics classrooms of urban schools. *Urban Education, 30*(4), 449–475.

Campbell, P.F. (2007). *The development, activity, pedagogical practices and impact of mathematics coaches in elementary and middle schools*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Campbell, P.F., & White, D.Y. (1997). Project IMPACT: Influencing and supporting teacher change in predominantly minority schools. In E. Fennema & B. S. Nelson (Eds.), *Mathematics teachers in transition* (pp. 309–355). Mahwah, NJ: Erlbaum.

Carlisle School Administration, The. (2005). Carlisle public schools potential new initiatives, submitted to the Carlisle finance committee. Retrieved on April 1, 2007 from http://www.carlisle.mec.edu/district/business/CSABudget.pdf.

Carpenter, T.P., Fennema, E., Peterson, P., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: an experimental study. *American Educational Research Journal, 26*(4), 499–531.

Chapin, S. (1994). Implementing reform in school mathematics. *Journal of Education, 176*(1), 67–76.

Clotfelter, C.T., Glennie, E., Ladd, H.F., & Vigdor, J. (2006). *Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina* (NBER, Working Paper No. 12285). Cambridge, MA: National Bureau of Economic Research.

Clotfelter, C.T., & Ladd, H. (1996). Recognizing and rewarding success in public schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education.* Washington, DC: The Brookings Institute.

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2007). *How and why do teacher credentials matter for student achievement* (NBER, Working Paper No. 12828). Cambridge, MA: National Bureau of Economic Research.

Coleman, J.S., Campbell, E.Q., Hobson, C.J. McPartland, J., Mood, A.M., Weinfeld, F.D., et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Department of Education.

Conant, J.B. (1963). *The education of American teachers*. New York: McGraw-Hill.

Cooper, S.T., & Cohn, E. (1997). Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review, 16*(3), 313–327.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence* (Document R-99-1). Seattle: Center for the Study of Teaching and Policy.

Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis, 23,* 57–77.

Darling-Hammond, L., Holtzman, D.J., Gatlin, S.J., & Heilig, J.V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13*(42). Retrieved on February 13, 2008 from http://epaa.asu.edu/epaa/v13n42/.

Davis, E.A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research, 76*(4), 607–651.

Decker, P.T., Mayer, D.P., & Glazerman, S. (2004). *The effects of Teach for America on students: Findings from a national evaluation.* Princeton, NJ: Mathematica Policy Research, Inc.

Dee, T.S., & Keys, B. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management, 23*(3), 471–88.

Dolton, P., & van der Klaauw, W. (1995). Leaving teaching in the UK: A duration analysis. *The Economic Journal, 105***,** 431–444.

Eberts, R., Hollenbeck, K., & Stone, J. (2002). Teacher performance incentives and student outcomes. *Journal of Human Resources, 37*(4), 913–927.

Eisenberg, T.A. (1977). Begle revisited: Teacher knowledge and student achievement in algebra. *Journal for Research in Mathematics Education, 8*(3), 216–222.

Feiman-Nemser, S., Schwille, S., Carver, C., & Yusko, B. (1999). *A conceptual review of literature on new teacher induction*. East Lansing, MI: National Partnership for Excellence and Accountability in Teaching, Michigan State University.

Fennell, F. (2006). We need elementary school mathematics specialists now. *NCTM News Bulletin*. *43*(4).

Ferrini-Mundy, J., & Johnson, L. (1996). Chapter 7: Highlights and implications. In J. Ferrini-Mundy & T. Schram (Eds.), *Journal for Research in Mathematics Education* (Monograph 8) (pp. 111-128, 129–133).

Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives*, *7*(9). Retrieved February 13, 2008, from http://epaa.asu.edu/epaa/v7n9.html.

Figlio, D.N., & Kenny, L.W. (2006). *Individual teacher incentives and student performance* (NBER, Working Paper No. 12627). Cambridge, MA: National Bureau of Economic Research.

Figlio, D.N., & Kenny, L.W. (2007). Individual teacher incentives and student performance. *Journal of Public Economics, 91*(5–6), 901–914.

Fulton, K., Yoon, I., & Lee, C. (2005). *Induction into learning communities*. Washington, DC: National Commission on Teaching and America’s Future.

Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

Gearhart, M., Saxe, G.B., Seltzer, M., Schlackman, J., Ching, C.C., Nasir, N., et al. (1999). Opportunities to learn fractions in elementary mathematics classrooms. *Journal for Research in Mathematics Education*, *30*(3), 286–315.

Glazerman, S., Senesky, S., Seftor, N., & Johnson, A. (2006). *Design of an impact evaluation of teacher induction programs: Final report*. Princeton, NJ: Mathematica Policy Research, Inc.

Glewwe, P., Ilias, N., & Kremer, M. (2003). *Teacher incentives* (NBER, Working Paper No. 9671). Cambridge, MA: National Bureau of Economic Research.

Goldhaber, D. (2007). *Everyone’s doing it, but what does teacher testing tell us about teacher effectiveness?* (Center for Analysis of Longitudinal Data in Education Research, working paper 9). Washington, DC: Urban Institute.

Goldhaber, D.D., & Brewer, D.J. (1997a). Evaluating the effect of teacher degree level on educational performance. *Developments in School Finance*, 197–210.

Goldhaber, D.D., & Brewer, D.J. (1997b). Why don’t schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *The Journal of Human Resources*, *XXXII*(3), 505–523.

Goldhaber, D.D., & Brewer, D.J. (2000). Does teacher certification matter? High school certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*, 129–146.

Goldhaber, D., DeArmond, M., Lui, A., & Player, D. (2007). *Returns to skill and teacher wage premiums: What can we learn by comparing the teacher and private sector labor markets?* (working paper 8). Seattle: School Finance Redesign Project.

Gordon, R., Kane, T.J., & Staiger, D.O. (2006). *Identifying effective teachers using performance on the job*. Washington, DC: The Brookings Institution.

Greenwald, R., Hedges, L., & Laine, R.D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*(3), 361–396.

Gruber, K.J., Wiley, S.D., Broughman, S.P., Strizek, G.A., & Burian-Fitzgerald, M. (2002). School and staffing survey, 1999-2000: An overview of the data for public, private, public charter, and bureau of India affairs elementary and secondary schools. Washington, DC: U.S. Department of Education. Retrieved on November 1, 2007 from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002313.

Hanushek, E. (1986). The economics of schooling—production and efficiency in public schools. *Journal of Economic Literature*, *24*(3), 1141–1178.

Hanushek, E. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, *18*(4), 45–51, 62.

Hanushek, E., Kain, J.F., & Rivkin, S.J. (2004). Why public schools lose teachers. *The Journal of Human Resources*, *39*(2), 326–354.

Harbison, R.W., & Hanushek, E.A. (1992). *Educational performance of the poor: Lessons from rural northeast Brazil* (pp. 81–177). Washington, DC: World Bank.

Harris, D.N., & Sass, T.R. (2007a). *The effects of NBPTS-certified teachers on student achievement.* (National Center for Analysis of Longitudinal Data in Education Research, working paper #4). Washington, DC: Urban Institute.

Harris, D.N., Sass, T.R. (2007b). *Teacher training, teacher quality and student achievement.* (National Center for Analysis of Longitudinal Data in Education Research: working paper #3). Washington, DC: Urban Institute.

Hawkins, E.F., Stancavage, F.B., & Dossey, J.A. (1998). *School Policies and Practices Affecting Instruction in Mathematics: Findings From the National Assessment of Educational Progress.* Washington, DC: National Center for Educational Statistics.

Hedges, L.V., Laine, R., & Greenwald, R. (1994). An Exchange: Part I: Does money matter? A meta-analysis of studies of the effects of differential school input on student outcomes. *Educational Researcher*, *23*(3), 5–14.

Hill, H.C., Ball, D.L., & Schilling, S.G. (in press). Unpacking “pedagogical content knowledge”: Conceptualizing and measuring teachers’ topic-specific knowledge of students. *Journal for Research in Mathematics Education.*

Hill, H.C., Rowan, B., & Ball, D.L. (2005). Effects of teachers’ mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*, 371–406.

Horowitz, S.H., Fuchs, L., & Clarke, B. (2006*). [Written comments to the national mathematics advisory panel.] 1-3*. Retrieved on September 1, 2007 from http://www.ed.gov/about/bdscomm/list/mathpanel/3rd-meeting/presentations/wendorf.james.public.pdf.

Inge, V. (2002, September 20). A Virginia forum: Moving from teacher leaders to mathematics teacher specialist: Notes from the Virginia mathematics & science coalition forum.

Ingersoll, R.M., (2000). Turnover among mathematics and science teachers in the U.S. Retrieved on October 1, 2007 from http://repository.upenn.edu/gse_pubs/96.

Ingersoll, R.M., & Kralik, J.M. (2004). *The impact of mentoring on teacher retention: What the research says*. Denver: Education Commission of the States.

Jacob, B.A., & Lefgren, L. (2002). The impact of teacher training on student achievement: Quasi-experimental evidence from school reform efforts in Chicago. *The Journal of Human Resources, 39*(1), 50–79.

Kagan, D.M. (1992). Professional growth among pre-service and beginning teachers. *Review of Educational Research, 62*(2), 129–169.

Kane, T.J., Rockoff, J.E., & Staiger, D.O. (2006). What does certification tell us about teacher effectiveness? Evidence from New York City. Retrieved on October 1, 2007 from http://www.dartmouth.edu/~dstaiger/Papers/nyc%20fellows%20march%202006.pdf.

Karges-Bone, L., Collins, F.J., & Maness, E. (2002). Last place is a fine place to begin. *Journal of Staff Development, 23*(2), 24–48.

Kemis, M., Heiting, T., Spitzli, J., & Lang, E. (2003). 2002 Iowa mathematics and science needs assessment: A study of Iowa teachers, superintendents, AEA mathematics and science coordinators, and higher education faculty. Research Institute for Studies in Education, Iowa State University: Iowa Department of Education.

King Rice, J. (2003). The incidence and impact of teacher professional development: Implications for education productivity. In M.L. Plecki, & D.H. Monk (Eds.), *School finance and teacher quality: Exploring the connections* (pp. 111–136). Larchmont, NY: Eye on Education, Inc.

Koehler, M.J., & Lehrer, R. (1998). Designing a hypermedia tool for learning about children's mathematical cognition. *Journal of Educational Computing Research, 18*(2), 123–145.

Krueger, A.B., (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*. *114,* 497–532.

Laczko-Kerr, I., & Berliner, D.C. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy. *Education Policy Analysis Archives, 10*(37). Retrieved July 25, 2007 from http://epaa.asu.edu/epaa/v10n37/.

Ladd, H. (1999). The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review, 18*(1), 1–16.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37–62.

Larson, J.C. (2000). The role of teacher background and preparation in students' algebra success. Rockville, MD: Office of Shared Accountability, Montgomery County Public Schools.

Lavy, V. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy 110*(6), 1286–1316.

Lavy, V. (2004). *Performance pay and teachers' effort, productivity and grading ethics*. Working paper 10622. Cambridge: National Bureau for Economic Research.

Learning Mathematics for Teaching Project (2006). *Technical materials on the number and operations content knowledge measure*. Ann Arbor, MI: Author.

Levine, A. (2006). *Educating school teachers*. Washington, DC: The Education Schools Project.

Loeb, S., & Page, M. (2000). Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and non-pecuniary variation. *Review of Economics and Statistics, 82*(3), 393–408.

Lopez, A., Lash, A., Schaffner, M., Shields, P., & Wagner, M. (2004). *Report of research on the impact of beginning teacher induction on teacher quality and retention*. Menlo Park, CA: SRI International.

Mandeville, G.K., & Liu, Q.D. (1997). The effect of teacher Certification and task level on mathematics achievement. *Teaching and Teacher Education, 13,* 397–407.

Marvel, J., Lyter, D.M., Peltola, P., Strizek, G.A., & Morton, B.A. (2006). Teacher attrition and mobility: Results from the 2004–05 teacher follow-up survey (NCES 2007-307). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Maryland State Department of Education (2001). *Keys to Math Success: A Report from the Maryland Mathematics Commission*. Baltimore: Author.

McDevitt, T.M., Troyer, R., Ambrosio, A.L., Heikkinen, H.W., & Warren, E. (1995). Evaluating prospective elementary teachers' understanding of science and mathematics in a model preservice program. *Journal of Research in Science Teaching, 32*(7), 749–775.

McGinnis, J.R., Parker, C., & Graeber, A.O. (2004). A cultural perspective of the induction of five reform-minded beginning mathematics and science teachers. *Journal of Research in Science Technology, 41,* 720–747.

McGrath, C., & Rust, J. (2002). Academic achievement and between-class transition time for self-contained and developmental upper-elementary classes. *Journal of Instructional Psychology, 29*(1), 40–43.

Miller, J.W., McKenna, M.C., & McKenna, B.A. (1998). A comparison of alternatively and traditionally prepared teachers. *Journal of Teacher Education, 49*(3), 165–177.

Miller B., Moon, J., & Elko, S. (2000). *Teacher leadership in mathematics and science: Casebook and facilitators guide*. Portsmouth, NH: Heinemann.

Monk, D.H. (1992). Education productivity research: An update and assessment of its role in education finance reform. *Educational Evaluation and Policy Analysis, 14*(4), 307–332.

Monk, D.H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13,* 125–145.

Monk, D.H., & King, J.A. (1994). Multilevel teacher resource effects on pupil performance in secondary mathematics and science: The case of teacher subject-matter preparation. In R.G. Ehrenberg (Ed.), *Choices and consequences: Contemporary policy issues in education* (pp. 29–58). Ithaca, NY: ILR Press.

Moon, J. (2002). Guest editorial: Why research has a role with teachers. *American Journal of Undergraduate Research, 1*(3).

Mullens, J.E., Murnane, R.J., & Willett, J. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review*, *40*, 139–157.

Muralidaran, K., & Sundararaman, V. (2006). *Teacher incentives in developing countries: Experimental evidence from India*. Boston: Department of Economics, Harvard University.

National Center for Research on Teacher Learning (1991). *Findings from the teacher education and learning to teach study: Final report.* East Lansing, MI: The National Center for Research on Teacher Learning.

National Commission on Mathematics and Science Teaching for the 21st Century. (2000). *Before it's too late: A report from the National Commission on Mathematics and Science Teaching for the 21st Century*. Washington, DC: U.S. Department of Education.

National Commission on Teaching and America's Future. (2003). *No dream denied: A pledge to America's children*. Washington, DC: Author.

National Council of Teachers of Mathematics (2000). P*rinciples and standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics (2002). *Induction and mentoring of new teachers*. Retrieved July 31, 2007, from www.nctm.org/about/content.aspx?id=6362.

National Research Council, Center for Education, Division of Behavioral and Social Sciences and Education (2001). *Adding it up: Helping children learn mathematics*. J. Kilpatrick, J. Swafford, & B. Findell (Eds.). Mathematics learning study committee. Washington DC: National Academy Press.

Noell, G.H. (2006). *Annual report of value added assessment of teacher preparation*. Baton Rouge, LA: Louisiana State University.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257

Ohanian, S. (1998). *Tiger in the children's garden: The Exxon education foundation K–3 mathematics specialist*. Dallas: Exxon Educational Foundation.

Podgursky, M., & Springer, M. (in press). Teacher performance pay. A review. *Journal of Policy Analysis and Management.*

Raymond, M., Fletcher, S., & Luque, J. (2001). *Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Stanford, CA: CREDO.

Reed, D., Rueben, K.S., & Barbour, E. (2006). *Retention of new teachers in California*. San Francisco: Public Policy Institute of California.

Reys, B.J., & Fennell, F. (2003). *Who should lead mathematics instruction at the elementary school level? A case for mathematics specialists*. Reston, VA: National Council of Teachers of Mathematics.

Richards, C.E., & Sheu, T.M. (1992). The South Carolina school incentive reward program: A policy analysis. *Economics of Education Review, 11,* 1, 71–86.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458.

Romney, M. (2007). *The governor's budget recommendation, house 2, fiscal year 2007: Education reform #10, section 50, chapter 74B.* Retrieved on August 13, 2007 from http://www.mass.gov/bb/fy2007h1/2007outsec/.

Rowan T.E., & Campbell, P.F. (1995). *School-based mathematics specialists providing on site support for instructional reform in urban mathematics classrooms. The project impact model.* Arlington, VA: National Science Foundation. Retrieved on November 1, 2007 from http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/8a/49.pdf.

Rowan, B., Chiang, F., & Miller, R.J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education, 70,* 256–284.

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.

Ruehl, S., & Wray, J. (2006). PreK–8 and 9–12 Maryland LEA mathematics specialist survey. Ellicott City, MD: Howard County Public School System.

Sanders, W., & Rivers, J. (1996) *Cumulative and residual effects of teachers on future student academic achievement.* Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Santibáñez, L., Martínez, J.F., Datar, A., McEwan, P.J., Setodji, C.M., & Basurto-Dávila, R. (2007). *Breaking ground: Analysis of the assessment system and impact of Mexico's teacher incentive program "carrera magisteria."* Santa Monica, CA: RAND Education.

Saxe, G.B., Gearhart, M., & Nasir, N.S. (2001). Enhancing students' understanding of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education, 4*(1), 55–79.

Sconiers, S. (1991). The UCSMP elementary mathematics specialist project. In I. Wirszup, & R. Streit (Eds.), Developments in school mathematics education around the world. *Teaching Children Mathematics, Vol. 2,* 58, 253–260.

Sheehan, D.S., Marcus, M. (1978). Teacher performance on the national teachers examination and student mathematics and vocabulary achievement. *The Journal of Educational Research, 71*(3), 134–136.

Slotnik, W., Smith, M.D., Glass, R.J., Helms, B.J., & Ingwerson, D.W. (2004). *Catalyst for change: Pay for performance in Denver. Final report*. Denver: Community Training and Assistance Center.

Strizek, G.A., Pittsonberger, J.L., Riordan, K.E., Lyter, D.M., & Orlofsky, G.F. (2006). Characteristics of schools, districts, teachers, principals, and school libraries in the United States: 2003–04 schools and staffing survey (NCES 2006-313 Revised). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Tatto, M.T., Nielsen, H.D., Cummings, W., Kularatha, N.G., & Dharmadasa, K.H. (1993). Comparing the effectiveness and costs of different approaches for educating primary school teachers in Sri Lanka. *Teaching and Teacher Education, 9*(1), 41–64.

Thompson, L. (2007). Math coaches help teachers help kids. *The Seattle Times.* Retrieved on February 12, 2008, from http://seattletimes.nwsource.com/html/education/2003713566_mathcoaches19m.html

Totterdell, M., Bubb, S., Woodroffe, L., & Hanrahan, K. (2004). *The impact of newly qualified teachers (NQT) induction programmes on the enhancement of teacher expertise, professional development, job satisfaction or retention rates: A systematic review of research literature on induction.* London, England: EPPI-Centre.

Van Haneghan, J.P., Pruet, S.A., & Bamberger. (2004). Mathematics reform in a minority community: Student outcomes. *Journal of Education for Students Placed at Risk*, *9*(2), 189–211.

Virginia General Assembly (2007). Summary of the amended 2006-2008 biennial budget adopted by the 2007 general assembly affecting the direct aid to public education and department of education central office budgets. Retrieved on May 10, 2007 from http://www.doe.virginia.gov/VDOE/suptsmemos/2007/reg001a.pdf.

Virginia Legislature (2005). Senate joint resolution no. 428, offered January 19, 2005. Requesting the board of education include the mathematics specialist endorsement in the licensure regulations for school personnel report. Retrieved on September 1, 2007 from http://leg1.state.va.us/cgi-bin/legp504.exe?051+ful+SJ428.

Virginia Legislature (2006). House joint resolution no. 258, offered February 2, 2006. Commending Virginia school boards that employ mathematics specialists in order to increase student achievement by increasing the quality of mathematics instruction. Retrieved September 1, 2007 from http://leg1.state.va.us/cgi-bin/legp504.exe?061+ful+HJ258.

Virginia Legislature (2007). House bill number 2093, offered January 10, 2007. A bill to amend and reenact §§ 22.1-253.13:1 through 22.1-253.13:7 and 22.1-274 of the code of Virginia, relating to the standards of quality.

Virginia Mathematics and Science Coalition, Special Issue: The role of the elementary mathematics specialists—overview. *Journal of Mathematics and Science*. (2002).

Wayne, A.J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*. *73*(1), 89–122.

West, L., & Staub, F.C. (2003). *Content-focused coaching*. Portsmouth, NH: Heinemann.

Winters, M., Ritter, G., Barnett, J., & Greene, J. (2006). *An evaluation of teacher performance pay in Arkansas.* Fayettville, AR: University of Arkansas, Department of Education Reform.

Wray, J. (2007). *PreK–8 and 9–12 mathematics leadership specialist certificate programs*. Unpublished Manuscript.

Wu, H. (2007). What should be taught in the elementary curriculum. Retrieved on May 1, 2007 from http://math.berkeley.edu/~wu/.