

The Teacher Preparation → Teacher Practices → Student Outcomes Relationship in Special Education: Missing Links and Next Steps

A Research Synthesis

November 2006

Laura Goe, Ph.D.
ETS



1100 17th Street N.W., Suite 500
Washington, DC 20036-4632
877-322-8700 • 202-223-6690
www.ncctq.org

Copyright © 2006 National Comprehensive Center for Teacher Quality, sponsored under government cooperative agreement number S283B050051. All rights reserved.

This work was originally produced in whole or in part by the National Comprehensive Center for Teacher Quality with funds from the U.S. Department of Education under cooperative agreement number S283B050051. The content does not necessarily reflect the position or policy of the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by the federal government.

The National Comprehensive Center for Teacher Quality is a collaborative effort of Education Commission of the States, ETS, Learning Point Associates, and Vanderbilt University. 1706_11/06



Contents

	Page
Abstract.....	1
Introduction.....	2
Methods.....	4
The Changing Landscape of Special Education Teacher Preparation.....	5
The Current State of Research.....	8
Scientific Evidence.....	8
Types of Research and How It Is Used.....	10
Outcome Measures and Accountability.....	11
Testing for Special-Needs Students.....	12
Measuring Individual Teacher Effectiveness.....	13
Challenges to Conducting Research on the Teacher Preparation→	
Teacher Practices→ Student Outcomes Relationship.....	15
Effect Size.....	15
Context.....	16
Differences in Selectivity Among Teacher Programs.....	18
Fidelity of Implementation.....	19
Research on <i>General</i> Teacher Preparation.....	21
Research on the Teacher Preparation→ Teacher Practices→ Student Outcomes	
Relationship for Students With Special Needs and their Teachers.....	23
A Look at Several Studies.....	23
Recommendations.....	28
1. Encourage and Support Research.....	28
2. Extend the Research to Include Induction and Early Professional Development.....	28
3. Consider Curriculum-Based Measures of Achievement for Students	
With Special Needs.....	29
4. Include Special Education Instruction in all Teacher Preparation Programs, and	
Encourage Collaboration Between General and Special Education Teachers.....	29
5. Design Research That Measures Collective Contributions to Student Outcomes.....	30
6. Develop a Comprehensive Longitudinal Teacher Database.....	30

Conclusion	31
References.....	32

Abstract

The original goal of this paper was to document and analyze the research on the connection between teachers' preparation to teach special education students, their instructional practices once in the classroom, and their students' eventual learning achievement (the teacher preparation→ teacher practices→ student outcomes relationship for students with special needs and their teachers). However, a search of the literature revealed that this three-part relationship has gone virtually unexplored, not only in special education but also in general education. Although a synthesis of existing research was thus found to be impossible, there is still much to be learned about this relationship. The search illuminated the practical and methodological challenges in doing such research, which led to a set of recommendations for how to develop appropriate studies in the future.

Introduction

This report examines the relationship among teacher preparation → teacher practices → student outcomes for students with special needs and their teachers. The framing of this relationship with arrows is intended to emphasize its ordered nature, with preparation leading to practices, which then result in student outcomes. This ordered relationship will be examined in its entirety because it is crucial to understanding what teacher preparation programs need to do in order to better equip special education teachers with practices that will lead to improvements in student outcomes. Such outcomes include achievement tests and other measures of importance to the families of these students and to the schools that serve them. Thus, this report provides specific suggestions to further the goal of using research and state policy to inform the special education teacher preparation curriculum.

For purposes of this report, *teacher preparation* means the preparation program (generally in a college of education) that a teacher attended to obtain certification. This includes programs in schools of education and alternative programs that lead to certification in special education. *Teacher practices* are the instructional strategies and classroom behaviors that teachers exhibit as they interact with students in educational settings. *Student outcomes* are evidence of learning as exhibited by the student on various types of assessments; including standardized tests, teacher observations, curriculum-based measures, and other mechanisms for evaluating student progress. The connections among these three topics are represented by arrows to clarify the directional relationship that exists among them (i.e., that what a teacher learns and does has an effect on student learning).

The teacher preparation → teacher practices → student outcomes relationship for students with special needs and their teachers takes into consideration both educational inputs and outcomes. Most teacher quality studies on general classroom teachers have focused solely on inputs, such as education, certification, and experience, rather than on what teachers actually *do* in the classroom (see, for example, Ferguson & Womack, 1993; Goldhaber & Brewer, 1999; Milanowski, 2004; Mullens, Murnane, & Willett, 1996; Sanders, Skonie-Hardin, Phelps, & Minnis, 1994). Other teacher quality research has defined teacher quality by outcomes (i.e., what *students actually learn* in the classroom), typically measured by standardized tests—a sort of backward mapping in which teacher quality is defined empirically by students' test scores (see, for example, Fetler, 1999; Monk, 1994; Rockoff, 2004; Vandevort, Amrein-Beardsley, & Berliner, 2004). Some of this research is methodologically quite sophisticated, such as the value-added model developed by William Sanders (see Sanders, Saxton, & Horn, 1997, for a description). But approaches based solely on inputs or outcomes fail to account for what teachers actually *do* in classrooms (i.e., the instructional practices they use, many of which were learned in teacher preparation programs). Furthermore, these input and outcomes studies seldom take into consideration the context in which teachers make their contributions to student learning. These contexts may include school climate, collegial and administrative support, curriculum, peer effects, parental influence, and other factors. This is primarily because such studies are done at arm's length—the analysis of decontextualized data is done by researchers who never step inside the schools or meet the teachers who are the source of the data.

The other part of the connection relating what teachers learn in their teacher preparation programs to what they do in the classroom is an area of great interest and great challenge to institutions of higher education (Wineburg, 2006). There is no standard way to measure the effectiveness of teacher preparation programs; programs currently develop their own strategies and implement their own effectiveness studies.

As will be shown in this report, the research based on the teacher preparation→ teacher practices→ student outcomes relationship is very thin. Moreover, where research exists, it is often not conducted in such a way that the results can be applied widely. The reasons for the paucity of high-quality research in this area will be discussed at length. Recommendations will be made for how to promote research that can be used to provide teacher preparation programs with information that they can use to shape the curriculum and content of programs to train all who teach students with special needs, *whether or not they are special education teachers*. There also will be a discussion of the current landscape of special education teacher preparation, which continually shifts in response to legislation and the growing understanding of the needs of students with special needs.

In summary, the original objective for this report was to develop a research synthesis that would shed light on the teacher preparation→ teacher practices→ student outcomes relationship for special education teachers and their students. However, literature searches revealed that no research makes this *complete* connection, though some research illuminates *parts* of the connection. The discovery of the gaps in the research led to an approach that (1) describes the challenges that impede progress in making the teacher preparation→ teacher practices→ student outcomes connection, (2) points out where (and how) the research would need to be focused and/or expanded to make the important connections in teacher preparation→ teacher practices→ student outcomes, (3) identifies and discusses some representative research that met part of the criteria, and (4) makes recommendations to address some of the issues identified in the literature analysis.

Methods

Searches for research that appeared to focus on the teacher preparation→ teacher practices→ student outcomes connection were conducted initially through a Web search, readings provided by experts in the field, and a hand search.

The Web search was conducted using various configurations of key terms: special education research, teacher practices, teacher preparation, student achievement, teacher effectiveness, student outcomes, and teacher qualifications. While many studies and reports were located using those terms, there were virtually no research studies that made the teacher preparation→ teacher practices→ student outcomes connection. Several researchers in the field of special education were asked for seminal works that might be relevant to this topic, and they provided a number of leads. However, none of the studies the researchers provided focused on the entire teacher preparation→ teacher practices→ student outcomes connection. These papers did provide useful reference lists and/or information about the challenges in doing special education research, and the theoretical and philosophical discussions in the field. The hand search was conducted by going through special education journals to identify papers that addressed the general topic of special education research and then following a trail of references to locate additional studies that might have relevant findings. From those studies as well as from those provided by the experts, additional trails were developed and followed in the effort to locate relevant studies. As with the Web search, no research studies were found through this mechanism. There were, however, many studies that made *part* of the connection, (i.e., connecting teacher preparation to teacher practices, or connecting teacher practices to student learning, or connecting teacher preparation to student learning).

Because of the dearth of studies that connected teacher preparation→ teacher practices→ student outcomes, the focus of this report was necessarily reconsidered and a different strategy was employed: that of selecting representative research studies, discussing their relationship to the teacher preparation→ teacher practices→ student outcomes connection, and describing how such research might be extended to comprise the entire connection.

The Changing Landscape of Special Education Teacher Preparation

During the past 30 or 40 years, there has been a major shift in the way students with special needs are educated. Formerly, students with high-incidence disabilities (such as reading disabilities) went unidentified and were educated in general classrooms. General education teachers did not always recognize these disabilities, and even if they did, they may not have had the tools to effectively help the students in the general classroom. Students with more profound learning disabilities were typically taught in isolated, self-contained classrooms by a single teacher for the entire day.

Two important changes have occurred. First, there is now an increased effort to identify and assist students with disabilities while keeping them in general education classrooms for most of their instructional needs. Second, students with more profound disabilities are frequently placed in general education classrooms for all or part of the day, sometimes with a paraprofessional assisting them, and sometimes with a special education teacher advising and assisting the teacher. Thus, the great majority of students receiving special education now spend at least some part of every day in nonspecial education classrooms with their age-mates. This change has been brought about by federal legislation and by a deeper understanding of how students with special needs learn. Teacher preparation also has been changing in response to the emphasis to educate special-needs students in ways that support their diverse needs—be they educational, emotional, behavioral, or cultural (see, for example, Kavale, 2005; Maheady, 1997; Pugach, 2005; Pugach & Seidl, 1995).

As noted earlier, one of the most important changes in teacher education is the advent of a variety of combined programs. These programs break down the barriers separating courses for those who will teach *only special education* and courses for those who will teach *only general education*. In a number of teacher preparation programs, teacher candidates graduate with certifications in both special education and general education. In the book *Teacher Education in Transition* (Blanton, Griffin, Winn, & Pugach, 1997), details are offered on such programs in a number of different colleges. This volume grew out of a project of the Council for Exceptional Children's Teacher Education Division, which developed case studies of 10 programs selected for their commitment to integrating the education of special education and general education teachers. These programs also were selected to represent a range of sizes, missions, and governance methods and a variety of geographic regions. Each of the 10 programs described in this volume has unique features, but all share a common theme of ensuring that all teachers are well prepared to teach a diverse group of students in a variety of settings. Griffin and Pugach (1997) concluded from these case studies of collaboration that successful programs exist, and that this success is dependent on a number of factors including: strong administrative leadership; strong school partnerships, including professional development schools; commitment to evaluation of the programs; effective communication strategies; and willingness to consider changing their collective vision of the fundamental nature of teaching and learning. In addition, Griffin and Pugach point to the programs' successes in working with state departments of education and with the communities in which the programs are located.

In their report titled *Special Education in an Era of School Reform: Preparing Special Education Teachers*, Hardman, McDonnell, and Welch (1998) provide recommendations for moving toward preparation that involves:

(a) collaboration and cross-disciplinary training, (b) a common core of knowledge and skills for both general and special education teachers, and (c) field-based training that involves building and sustaining partnerships between higher education and the public schools. (p. 2)

The interpretation and application of these goals has been different for each program, however, as evidenced by the several different lists the authors provide (from such sources as ETS, the Joseph P. Kennedy Jr. Foundation, the U.S. Department of Education, and the University of Wisconsin–Milwaukee) describing what constitutes a common core of learning for programs preparing special education teachers. These differences in emphasis, philosophy, and structure representing the variability among teacher education programs need to be explored and connected to student outcomes.

Greater collaboration between special and general education teachers has helped bring increased attention to the importance of identifying struggling students and providing appropriate interventions to prevent their placement in special education. In the 2004 reauthorization of the Individuals with Disabilities Education Act (IDEA), emphasis is placed on providing early support for students who are struggling and, therefore, schools are expected to develop mechanisms to identify and help these students. One such mechanism is referred to as Response to Interventions (RTI), the purpose of which is not only to identify students in need of special education services but also to identify students in need of assistance *before* they are officially referred for special education. In a sense, the RTI is a treatment since the students' responsiveness to a well-designed and implemented research-based intervention helps determine whether the student is a candidate for more specialized services.

A 2006 white paper by the National Association of State Directors of Special Education and the Council of Administrators of Special Education, two important groups that help shape special education policy, put it this way:

One of the inherent problems with current special education programming as identified by the President's Commission and by Congress in IDEA 2004 is that the current system uses a wait to fail model before interventions and services are provided. Not surprisingly, both the President's Commission and Congress called for early intervention that addresses learning and behavioral issues rather than waiting for children to fail before intervention occurs. (p. 4)

The paper specifically refers to ensuring that all instructional staff is prepared to assist in such interventions: "Successful implementation of RtI (sic) depends on the ability of general and special educators to use RtI reliably and validly" (p. 6).

The importance of RTI to teacher preparation lies in the way it has changed the preparation of general and special education teachers and their practices. In previous years, classroom teachers

were likely to document a student's struggles and then make a referral to a special education teacher for evaluation. In most states and districts, the regular classroom teacher's instructional responsibilities for the student ended once he or she was transferred to the special education teacher. With the advent of RTI, however, a more formal collaborative process is needed. Both general and special education teachers may need training, either through their teacher preparation programs or through targeted professional development, in how to work together to decide upon an appropriate research-based intervention. Developing a strategy for implementation and evaluation also is necessary, and both special education and general education teachers need training in how to develop such plans and assess results.

For the purposes of considering the teacher preparation → teacher practices → student outcomes relationship for students with special needs and their teachers, RTI increases the difficulty of sorting out whose practices are responsible for student outcomes—the special education teacher or the general education teacher. The way in which RTI complicates attribution of student outcomes is a good example of how difficult it is to evaluate an *individual* teacher's contribution to student learning, in isolation from the contributions of other teachers or instructors (such as tutors or paraprofessionals). Should the general education teacher take sole credit for improved achievement as a result of an intervention that was designed and delivered collaboratively? And if both teachers should receive appropriate credit, how does a researcher sort out how much credit each should receive? If a tutor was employed to assist the student or the student was sent to a reading recovery teacher, how does that influence the crediting of improved student outcomes to the general education teacher? This is a dilemma that occurs not only in trying to evaluate the teacher preparation → teacher practices → student outcomes relationship for special education teachers but also for general education teachers who share the responsibility for student learning with other teachers, through team teaching or teaching across the curriculum (i.e., teaching reading and writing in every subject). Attributing student outcomes to individual teachers is quite a complex process, not simply a matter of identifying the teacher who is responsible for teaching a particular subject.

The Current State of Research

Before examining the state of research on special education teachers, the current *overall* state of research regarding teacher preparation→ teacher practices→ student outcomes relationship will be considered. This relationship is complicated by several factors. For example, how do we know that what is taught in teacher preparation programs actually translates into teacher practices that help students learn? Logically, we would have to be able to trace course content to a specific classroom practice, then connect that specific practice to measurable learning outcomes (such as a score on a test). Although this sounds simple in principle, the actual process is fraught with practical difficulties and therefore fragmented, so that only portions of the relationship are reported on individually in the literature. When we consider the additional requirement of looking specifically at special education teachers, we are complicating the picture even further. Perhaps more than any other program area, special education teacher preparation is in a state of flux, primarily because of changing beliefs about how best to train teachers to work with students with special needs. The variability in types of programs is substantial, and what teachers know and believe when they graduate from a special education teacher preparation program may differ greatly depending on where they have completed their degree coursework for initial certification.

Another factor that complicates research on the teacher preparation→ teacher practices→ student outcomes relationship is that special education teachers are no longer exclusively responsible for the outcomes of students with special needs. General education teachers have increasingly been required to take significant roles in educating students with special needs, and some have had little or no coursework or training in working with these students. When research regarding the outcomes of students with special needs is considered, it must now be understood that there may be many teachers contributing to those outcomes (particularly in the secondary grades). However, there has been substantial movement toward preparing teachers simultaneously for general and special education.

The desired result of this movement is that even those teachers who do not seek certification or endorsement in special education will be very well prepared to work with students with special needs. This means that when conducting research into teachers' contributions to the learning of special-needs students, it may be particularly important also to consider the contributions of the general education teachers trained to be effective in instructing special-needs students in the general classroom.

Scientific Evidence

Scientific evidence plays an important role in determining the most appropriate instructional practices for special education (Carnine, 1999). Since the advent of No Child Left Behind (NCLB), the push toward scientific evidence has led many researchers to focus on designing and conducting studies that meet high standards as defined by the What Works Clearinghouse (www.whatworks.ed.gov/). Research that is acceptable for the What Works Clearinghouse must meet the following criteria (What Works Clearinghouse, 2003):

- Employs systematic, empirical methods that draw on observation or experiment; involves data analyses that are adequate to support the general findings; relies on measurements or observational methods that provide reliable data; makes claims of causal relationships only in random-assignment experiments or other designs (to the extent such designs substantially eliminate plausible competing explanations for the obtained results).
- Ensures that studies and methods are presented in sufficient detail and clarity to allow for replication or, at a minimum, to offer the opportunity to build systematically on the findings of the research.
- Obtains acceptance by a peer-reviewed journal or approval by a panel of independent experts through a comparably rigorous, objective, and scientific review.
- Uses research designs and methods appropriate to the research question posed.

Thus, both qualitative and quantitative methods are deemed appropriate as tools to investigate effectiveness, and some researchers explicitly state the need for using *both* methods (e.g., Wilson, Floden, & Ferrini-Mundy, 2001).

Some researchers believe that more research should be done using experimental designs (see, for example, Cook, 2002; Slavin, 2004). Randomized studies are considered the gold standard of research seeking causal explanations not just in education but in most fields (Riehl, 2006). Randomized experimental designs involve randomly assigning students, teachers, or schools to treatment or control groups. Such randomized studies support causal arguments because each teacher or student has an equal chance of being assigned to either group, and with large numbers of teachers or students involved in the study, there is no need to specify or to control for preexisting differences (for discussion of research design, see Shadish, Cook, & Campbell, 2002). Researchers typically focus on randomizing students primarily because the number of participants needed to make causal claims is large, and it is often difficult to obtain sufficient numbers of teachers to use them as the unit of randomization. However, such designs may overestimate effects because they fail to take into account the clustering of students in classrooms and schools (Bryk & Raudenbush, 1988; Schochet, 2005).

The difficulties in doing randomized experimental studies in education are many, including gaining cooperation from students, parents, administrators, and districts to accept random assignment to a control or treatment group—particularly when the treatment appears to be either particularly onerous or attractive. It is, therefore, not surprising that there have been comparatively few studies of this type in the field of education and none at all in the area that is the focus of this research synthesis. However, as noted in the What Works Clearinghouse definition, high-quality research can be conducted without randomized experimental designs, with the caveat that researchers may not make claims about causality unless they have used a randomized experimental design.

Ongoing discussions about research methodologies and how they matter in the study of education are common occurrences in the literature since the advent of NCLB (see, for example, Cook, 2002; Eisenhart & Towne, 2003; Raudenbush, 2005; Riehl, 2006; Slavin, 2004). The consensus among many researchers appears to be that it is more important to prioritize excellent

research design that is appropriate to the population and the circumstances and that uses the most rigorous methods available given those limitations.

Types of Research and How They Are Used

Gersten, Baker, and Pugach (2001) wrote a chapter called “Contemporary Research on Special Education Teaching” in the *Handbook on Research in Teaching*. In this chapter, they describe the state of special education research at the end of the 20th century. While this is a helpful work that takes into account many of the forces that influence and impede research in this area, there is little discussion of different types of research—how the research gets done is not a focus of the chapter. All research is treated as essentially equivalent. One of the goals of this report is to delve deeper into the different approaches to special education research and to explore the relative advantages and disadvantages of each.

In their article “Research in Special Education: Scientific Methods and Evidence-Based Practices,” Odom, Brantlinger, Gersten, Horner, Thompson, and Harris (2005) make a case that different research methodologies are needed because of the complexity of special education. They also note that effective special education practices have been identified using multiple methodologies.

Both qualitative and quantitative research methods are typically used to assess parts of the teacher preparation → teacher practices → student outcomes relationship in education for both general education and special education teachers. Qualitative methods typically involve the analysis of language-based data, including case studies, interviews, classroom observations, documents, artifacts, and open-ended survey questions. Quantitative methods involve the analysis of numerical data from surveys, classroom observations using a numeric rubric, and the collection and analysis of data on student achievement, often linked with specific teacher characteristics or specific schools.

Qualitative research methods are most useful for developing a deeper understanding of teacher practices in the field. They are able to help researchers better understand the *how* rather than just the *how many*. But because qualitative data are considerably more expensive than quantitative data to collect, qualitative studies are frequently limited to small numbers of participants. Another drawback to qualitative methods is that they are typically more intrusive than quantitative methods. Some teachers may be unwilling to participate because they are uncomfortable being observed or interviewed. Thus, locating teachers who are willing to have researchers visit their classrooms may require considerable effort and expense, and in the end, the group of willing teachers might not be truly representative, thus limiting the generalizability of the findings.

Quantitative studies usually involve many more cases, and the findings are typically expressed in statistical terms. Because of the greater numbers of teachers involved, it is generally possible to extrapolate findings from carefully performed analyses to other teachers. This is more difficult to do with qualitative research. Quantitative methods are typically considered more rigorous and the results more readily generalized to other populations (Johnson, 2001).

Outcome Measures and Accountability

The 1983 report *A Nation at Risk* helped initiate the push toward thinking about student achievement as it relates to teacher effectiveness (National Commission on Excellence in Education). Since then, student achievement has been increasingly emphasized by federal legislation such as NCLB. Even before federal legislation mandated student testing, states had begun tying accountability measures with student test scores at the school level (Goertz & Duffy, 2000; Goertz, 2000; Ladd, 2001). And as statistical methods and software have increased in sophistication, researchers have focused on tying student achievement to schools (Burstein, 1980; Cronbach, Linn, Brennan, & Haertel, 1997), classrooms (Bryk & Raudenbush, 1988), and ultimately, to individual teachers (Mendro, Jordan, Gomez, Anderson, & Bembry, 1998).

In recent years, researchers have increasingly been focusing on how teacher qualifications, characteristics, behaviors, and classroom practices are related to student outcomes. There has been an increased interest in students' performance on standardized tests as an outcome measure. As teachers have become recognized as the most important resource in student learning (see, for example, Darling-Hammond, 2000; Fetler, 1999; Hanushek, Kain, & Rivkin, 1998; Nye, Konstantopoulos, & Hedges, 2004; Rowan, Correnti, & Miller, 2002), teacher preparation has come under increasing scrutiny for its role in preparing a workforce of highly trained teachers (American Federation of Teachers, 2000; Allen & Palaich, 2000; The Teaching Commission, 2004; Wilson et al. 2001). Furthermore, the inclusion of more students, particularly students with special needs, in standardized testing (as required the reauthorization of NCLB) has spotlighted the role of teacher preparation programs in ensuring that those students with special needs are receiving high-quality instruction.

Expectations for students with special needs have changed considerably over the years as well, from the early days of their near-universal exclusion from the curriculum that other students were expected to master to the current emphasis on all students working toward high grade-level standards. This change in expectations for students has had the impact of changing expectations for teacher effectiveness. It is no longer enough for teachers to succeed in facilitating their students' attainment of individualized education program (IEP) goals; teachers also must prepare students for their place in the school accountability system.

Yet, despite the increased attention given to teacher effectiveness, there is little in the way of evidence, scientific or otherwise, that has convincingly clarified what teachers should be learning in their teacher preparation programs (Allen, 2003; Cochran-Smith & Zeichner, 2005). Nor is there solid agreement in the field about what teachers should be learning in order to increase their effectiveness (Shulman, 2005). There is a general push in teacher education to try to give teachers the appropriate skills and knowledge so that they are doing as follows:

- Creating the best possible opportunities for *all* students (including students with special needs) to achieve at high levels.
- Supporting students with special needs in their efforts to work at grade level.
- Using age-appropriate general education standards and curriculum for *all* students, including special-needs students.

Given the lack of definitive research about the relationship between what teachers learn in their preparation programs and how their learning influences what they do in classrooms and student outcomes, it is quite difficult to pinpoint the specific practices, approaches, and methods teachers should be using to provide *all* students with the best opportunities for learning.

Testing for Special-Needs Students

In the recent past, teachers, parents, and other school personnel worked together to decide upon appropriate expectations for individual students based on their current levels of achievement and their perceived ability to make progress. This resulted in an IEP that outlined specific, measurable goals for student learning and behavior. Meeting the goals of the IEP drove instruction. Outcomes, then, were not measured by standardized tests but rather by teacher-developed or other specialized, individually administered tests to determine progress toward goals, or by teacher observation and assessment of student performance. Although IEPs are still driving instruction for students with special needs, most students are no longer routinely exempted from taking standardized achievement tests with their general education peers. Today, schools include students with special needs in testing even if the results of the tests are not considered in calculating a school's progress toward meeting AYP goals, as required by NCLB. Thus, both general education and special education teachers are preparing students for the types of material they will encounter on these tests.

Many teachers believe that students with special needs should not participate in standardized testing because it may be a frustrating experience for those students, even with appropriate modifications of testing materials (Belden Russonello & Stewart Research and Communications, 2003). However, there are also experts in the field of testing policy who point out that under accountability pressures, schools will primarily pay attention to the students who are tested. This means that excluding any group (such as students with special needs) may result in fewer resources being devoted to the needs of that group because their scores will not affect the school's progress in meeting accountability goals (Heubert & Hauser, 1999; Linn, 2000). Moreover, some researchers have suggested that there are positive consequences to including students with disabilities in testing, such as "increased participation of students with disabilities in testing programs, higher expectations and standards, improved instruction, and improved performance" (Ysseldyke, Dennison, & Nelson, 2004, p. 1). But despite some arguments in favor of testing special-needs students, there are still troubling questions about how accurate such test scores are when different types of disabilities are taken into account (McDonnell, McLaughlin, & Morison, 1997).

Test score variance is only moderately affected by what teachers do in the classroom (Koretz, 2000) primarily because many standardized test scores are not necessarily sensitive to changes in instructional practice, particularly in the short term. This is partially a result of the way standardized tests are constructed and partially a result of the fact that teachers' abilities to influence student learning in particular subject areas are limited by the amount of time they spend with students between testing cycles. In a study of whether changes in instructional practice in secondary mathematics classes could affect student achievement, researchers' findings were mixed: The changes in practice made a difference in students' scores only when such changes were accompanied by changes in the curriculum (McCaffrey, Hamilton, Stecher,

Klein, Bugliari, & Robyn, 2001). This suggests that for at least some classroom practices, the impact on student learning is weak unless the practices are specifically designed to support particular curricular goals. Thus, linking particular teacher practices to student achievement is challenging given the small student effects that are likely to be observable in normal instructional interactions.

Measuring Individual Teacher Effectiveness

Standardized tests have become the ultimate measures of student learning; the measures that are privileged by the federal government in its NCLB policies. Standardized achievement tests have been used to measure teacher effectiveness in conjunction with such mechanisms as value-added models (Sanders & Horn, 1995, 1998; Webster & Mendro, 1997). The use of value-added models to measure teacher effectiveness has been criticized for a number of reasons, including their inability to explain the differences among teachers who have been designated as effective or ineffective (Kupermintz, 2003), and their exclusive reliance on standardized achievement scores (Kirby, McCaffrey, Lockwood, McCombs, Naftel, & Barney, 2002). However, such measures are often characterized as more objective than other approaches to measuring teacher performance (Teachers for a New Era, 2006).

A number of school districts, including Dallas and Denver, are experimenting with some type of value-added model in an attempt to identify and reward effective teaching. In California, value-added models are being considered by policymakers and researchers (Doran & Izumi, 2004). And in Tennessee where the concept of value added as a measure of teacher effectiveness first began, school administrators may consider teachers' value-added scores as part of the performance evaluation process. Ohio and Pennsylvania are both experimenting with value-added models. North Carolina and Michigan also are considering how best to use their student test data with value-added models.

Most value-added models require a minimum of two to three years worth of data to satisfy the model's statistical requirements. Most states, however, do not currently have the data systems to allow them to link individual teachers and students longitudinally (Data Quality Campaign, 2005). Thus, measuring the effectiveness of individual teachers through such mechanisms is not always an option.

There also have been recent attempts to measure the impact of teacher preparation programs on student learning by using value-added measures. Researchers in Louisiana have recently reported the results of using value-added measures to link teacher preparation programs to student achievement (Noell, 2006). The researchers found that there was a great deal of imprecision in the estimates of teacher effects, with large standard errors making it difficult to interpret their results.

Using standard growth models to measure teacher effectiveness is another potential mechanism for using tests to measure students' progress, holding constant variables such as students' poverty, race, language fluency, and prior test scores (Gong, 2004). But, once again, not every state or district has the data to perform such calculations. Typically, districts and states use some

combination of administrator observations and portfolios to try to measure individual teacher effectiveness.

States and districts are much better equipped to measure student performance at the school level by aggregating students' scores into a school-level score that can be used to measure adequate yearly progress (AYP) as required by NCLB. Although such scores are useful in comparing a school's performance within a particular district or state, they are less useful for establishing teacher effectiveness because the scores are not attributed to particular teachers but to the school as a whole.

Challenges to Conducting Research on the Teacher Preparation→ Teacher Practices→ Student Outcomes Relationship

One of the primary challenges in developing a clear picture of the teacher preparation→ teacher practices→ student outcomes relationship is that very few teacher preparation programs track their graduates after they leave the program. Thus, any research in this area must begin with the creation of a tracking system that will allow colleges that train teachers to keep track of their graduates over time. It is not surprising, then, that so little research has been done on the effectiveness of teacher education. Few teacher education programs have the resources to track where their graduates go over time and thus cannot recruit them to participate in any type of research. Wineburg (2006) found that 50 percent of the institutions she surveyed track students for about one year. Without mechanisms for continuous tracking, research on effectiveness in the classroom is severely hampered.

There are many possible explanations why teacher preparation programs do not keep track of their graduates. One likely explanation is that many states with a strong teachers union presence have great difficulty collecting teacher performance data at the individual teacher level due to union concerns about privacy and fears that such information may be used for purposes other than research. In many states, no longitudinal data on teachers is being collected. This adds another wrinkle to collecting data on teacher effectiveness. This fact became apparent when states were asked to submit equity plans in response to federal NCLB highly qualified teacher regulations requiring that all students, particularly those in high-poverty schools or schools not meeting their AYP goals, have an equal opportunity to be taught by a highly qualified, experienced teacher. The majority of states indicated that they did not have data systems in place that would permit them to compare teachers across schools¹. Furthermore, while some states have longitudinal data that enable them to track teachers, few states track which preparation program their teachers attended.

One state that does have data on teacher preparation programs and initial placements is California, a state that mandates induction programs for all beginning teachers. As part of this mandate, teachers must register in an online database for participation in an induction program. During this registration, the name of their schools and districts are collected as well as the college or university where they received their preparation for teaching. This type of data would help researchers locate teachers.

Effect Size

Another challenge in conducting research on the teacher preparation→ teacher practices→ student outcomes relationship lies in attempting to aggregate findings from studies that may use very different measures and methodologies. While no studies of this type were identified in special education, this is an ongoing issue in synthesizing research. Effect size estimates make it possible to compare findings from studies that used different methods and measures, particularly since few of these studies include effect sizes, making comparisons problematic (Cooper & Hedges, 1994). Effect sizes are becoming more prevalent in educational research but are hardly universal, particularly in qualitative research (for a discussion of effect sizes, see Grissom & Kim, 2005).

¹ See the plans at <http://www.ed.gov/programs/teacherqual/hqtplans/index.html>.

A common way to express effect size is Cohen's d , where the difference of the means of the treated and control groups are standardized by dividing them by their pooled variance (Cohen & Cohen, 1983). This transformation results in an estimate of the magnitude of the effect in terms of standard deviations which allows one to make meaningful comparisons of the size of different studies' results. In education research, effect sizes of interventions on student learning vary considerably, usually ranging from around .10 standard deviations to .70 standard deviations. Cohen (1988) points out that researchers often make the mistake of concluding that if a finding is statistically significant, it must be important. What is really crucial, he notes, is the size of the effects. His rule of thumb, accepted by many researchers today, defines effect sizes as: small, $d = .2$; medium, $d = .5$; and large, $d = .8$. Hattie (1992) conducted a meta-analysis of school interventions and discovered that simply spending a year in school has an effect size of around .40 on student learning (although that number may be somewhat inflated because it is difficult to take account of the length and intensity of the various interventions that were compared). Unfortunately, the research base does not yet exist to help us interpret effect sizes for teacher practices.

Effect sizes are particularly useful for comparing research studies because they allow us to examine multiple findings on a common scale. Authors of research studies frequently report the effect size, but when they do not, it can be approximated from the authors' reported test statistics (Grissom & Kim, 2005). Using effect sizes from various studies (either provided by the authors or calculated after the fact), a researcher developing a quantitative research synthesis can consider a key question: "Does the treatment help?" (Cooper & Hedges, 1994). Thus, if we had large numbers of studies that examined the teacher preparation → teacher practices → student outcomes relationship for students with special needs and their teachers, we could compare the *treatment*—the teacher preparation—and determine which teacher preparation programs resulted in practices that could be connected to better student outcomes. Unfortunately at this time, there is insufficient research to conduct such a quantitative synthesis.

Context

Another challenge that complicates research in this area is how to control for differences in context, including wide variability in the teachers' initial placements and the availability of and their participation in support structures such as induction programs or mentoring.

Some teacher preparation programs serve as *feeder programs* to certain local districts, meaning that many teachers in a given school share similar training and educational backgrounds. Such clustering of teachers within schools may matter in terms of research on teacher preparation and practices. The effects of being in a particular program may be enhanced by being in a school with other teachers who reinforce certain methods and philosophies. However, if teachers from particular preparation programs are scattered among schools that are highly variable in terms of context—such as percentage of students living in poverty—then it is difficult to sort out which factors are contributing to student outcomes. In fact, many context variables could confound the interpretation of the teacher preparation → teacher practices → student outcomes relationship.

The contexts within which teachers work can confound the impact of teacher preparation and ongoing teacher learning because contexts can make teachers more or less able to apply what

they learned to their particular situation. In a study of the California Formative Assessment and Support System induction program, Goe, Thompson, Paek, and Ponte (2005) found that in schools with high percentages of African-American students or low schoolwide performance as reflected in California's Academic Performance Index, teachers were less likely to have a high-engagement experience with the induction program. In another study, teachers who had high engagement with the induction program had students who outperformed the students of teachers with low engagement (Thompson, Paek, Goe, & Ponte, 2004). Although the mechanism that is responsible for this finding is not clear, it is possible that a challenging school context may distract teachers from engaging in new learning that is designed to build upon what teachers experienced as part of their teacher preparation programs. It also is possible that context variables influence teachers' decisions about their teaching in ways that affect student learning.

Some researchers have studied context and how it impacts studies of teachers and students (Achinstein & Barrett, 2004; McAllister, 2000; McLaughlin, Talbert, Bascia, & Stanford University Center for Research on the Context of Secondary School Teaching, 1990; Schofield & Sagar, 1979; Scribner, 1999; Sanders, Wright, & Horn, 1997). Contextual variables that might affect teachers' practices and contribute to differences in student learning outcomes include the following:

- The prescribed curriculum and its alignment with the texts and other instructional materials available to the teacher.
- Assistance from a curriculum supervisor in optimizing instructional practices that align with the curriculum.
- The availability of high-quality professional development that furthers teachers' content knowledge and/or helps them hone in on instructional practices that will increase students' learning (Cohen & Hill, 2000).
- The availability of support mechanisms such as trained mentor teachers and/or high-quality induction programs (for beginning teachers).
- Release time to work with mentor teachers, observe in experienced teachers' classrooms, meet with colleagues in grade or subject teams, or to participate in activities sponsored by the induction program (for beginning teachers).
- Physical facilities that are safe, comfortable, and conducive to instructional practices such as group work and coteaching (Corcoran, Walker, & White, 1988).
- School reform efforts that are supportive of teachers' instructional practice, including whole-school reforms and more isolated curricular reforms.
- Integration of special and general education (i.e., the degree to which special education and general education teachers work collaboratively to maximize the learning of all students, particularly those with special needs).
- School leadership (e.g., whether the school has effective leadership that supports teachers' efforts and minimizes distractions that would jeopardize the instructional time available) (Brewster & Klump, 2005).

- School climate and culture, including collegiality, support for teachers, treating teachers as professionals, and the ways in which teachers work together to share information about best practices and effective instructional techniques (Shann, 1998).
- Parental and community support for instruction and student learning, including providing students with encouragement and opportunities to bring learning into the home.
- Student engagement with learning and relationship with the educational system, (e.g., their attendance, cooperativeness, interest in learning)

There are many other context considerations that could be mentioned. The bottom line, however, is that there are many ways in which teachers' classroom practices are impacted, whether positively or negatively, by the setting in which they carry out instruction.

Differences in Selectivity Among Teacher Programs

Another reason it is difficult to conduct research on the teacher preparation → teacher practices → student outcomes relationship is because teacher education programs are not all alike in terms of their standards for admission. There are great differences among states in the minimum requirements to be admitted into a teacher preparation program. Some states require a bachelor's degree before a person can even enter a fifth-year teacher preparation program. In *The Making of a Teacher*, Feistritzer (1999) reports that 78 percent of institutions surveyed offer a four-year program while 11 percent require five years of study prior to initial certification. There also are differences among postbaccalaureate programs: 47.4 percent offer a postbaccalaureate program that leads to certification *but not a graduate degree*, and 43 percent offer certification *and a graduate degree*.

Within states, colleges may vary substantially in requirements for admission to their programs even if the programs themselves are similar. For example, admission to a University of California institution for teacher preparation is much more competitive and has more rigorous admission requirements than admission to a California State University. Despite this disparity, the eventual degree and certification will appear to be the same.

Thus, to compare graduates of different programs meaningfully, it is necessary to control for initial differences among teacher candidates. College admissions tests like the SAT or ACT may be useful in determining initial differences *before* college-level work was done. However, the differences in the quality of education received *at* the undergraduate level may mean that these initial test scores are no longer useful as a means of controlling for initial differences at the time the teacher candidates enter their preparation programs. If we assume, however, that degrees from elite institutions with challenging curricula represent greater learning than do degrees from less demanding colleges, then we also must concede that a person entering a teacher preparation program with a bachelor's degree in education has had different opportunities to learn than someone without a bachelor's degree. Grades may not adequately reflect differences in learning. For example, a 3.8 grade point average may be much harder to attain in one university than in another or in one field of study than another—even within the same university.

There is currently an increasing focus on measuring the learning of students as they graduate from college, both to determine what they have learned and whether they are ready to use their learning. A report by Dwyer, Millett, and Payne (2006) discusses the need for creating valid and reliable assessments that will measure the following:

- Workplace readiness and general skills.
- Domain-specific knowledge and skills.
- Soft skills such as teamwork, communication, and creativity.
- Student engagement with learning.

With the results of such assessments, researchers would have a better indication of what college graduates know at the time they enter teacher preparation programs (in the case of states that require a fifth year) or what they know when they join the ranks of certified teachers (in the case of states that permit students to complete their teacher preparation as part of their undergraduate training). Such measures serve the same function as pretests, which are commonly used to help determine the effectiveness of interventions. Until such assessments are in place, it will be difficult to control for differences among teachers as they begin their teacher preparation programs.

Fidelity of Implementation

Another serious issue that increases the difficulty of evaluating teacher practices and their impact on learning is the fidelity of the treatment. For example, if all teacher candidates in a teacher preparation program (1) learn the same best practices in their program, and (2) implement those practices in the best possible way under similar circumstances, then it would be a straightforward matter to measure the impact of teaching such practices. It is more likely, however, that teacher candidates do the following:

- Learn about many or most of the best practices, but have different interpretations of those practices and different ideas about how and when to use them.
- Implement the practices with wide variability in terms of how closely their use of the practice aligns with what was intended.
- Use the practices under widely differing circumstances (e.g., in high- versus low-performing schools; with minority versus nonminority students; with fluent speakers of English versus English language learners; or in resource, mainstreamed, sheltered, or all-day special education classrooms.)

Thus, teachers may report they are using a best practice learned in their teacher preparation program, but the way the practice is actually being used may be quite different from the way a cohort member in a different school is using the same practice.

Determining fidelity of implementation is difficult with data obtained from self-reports (i.e., surveys or interviews in which teachers are asked about their own practices). For a more accurate assessment of whether a teacher is using the practice as intended, a trained observer would go to a teacher's classroom, watch for occurrences of the practice, and then make judgments about

how closely the teacher's use of the practice follows what is intended. Consider the following example. Teachers learn various questioning strategies in their teacher preparation program. When asked if they use *wait time* when questioning students in whole-class discussions, they reply that they practice this strategy nearly all the time. A trained observer then observes their questioning strategies in several class sessions and records a teacher's use of wait time, along with the frequency of use, the circumstances in which it is used, and the length of each incident of wait time. From this data, the researcher can determine how close a teacher's implementation of the strategy is to what has been reported to be effective.

Unfortunately, this type of qualitative research is costly and time-consuming because of the person-hours involved in gathering and interpreting the data. However, it is the mechanism most likely to identify whether specific teaching practices are linked to specific student outcomes such as achievement. For example, if researchers gather information on a sample of teachers about their classroom wait time practices and then rank the teachers according to how well they were implementing the strategy, they can meaningfully determine the correlation between high- and low-fidelity practices and student outcomes.

It is not simple, of course, to isolate specific practices and to consider their relationship to outcomes. This is because a teacher who implements one questioning strategy at a high level may also be concurrently implementing other best practices with great fidelity. With the various practices involved in classroom instruction confounded in this way, it is difficult to sort out the effects of any one practice. In addition, even a best practice that is well implemented may not have enough of an impact on student learning to produce measurable achievement differences among classrooms.

Looking at a set of best practices, such as a constellation of questioning strategies that has been linked to better student learning outcomes, might be a way to raise the power of the strategies and increase the likelihood of seeing significant differences in student achievement among classrooms with high and low implementation. In other words, a *set* of practices might be more likely to have an impact on student achievement than a single practice.

Furthermore, there is concern with the authenticity of the strategies; since they generally do not occur in isolation, it makes little sense to study them in isolation. There is an inherent complexity in teaching that researchers should try to capture and analyze. However, this would greatly increase the research expenses and timeframe. The researcher would be observing, for example, a number of different questioning strategies and would have to find some way to appropriately weigh the teacher's score on each strategy to achieve some sort of total score. The total would be used to rank the teacher's degree of fidelity to the best questioning practices.

Even with qualitative research, quantitative research methods would have to be employed to control for differences in student ability and other factors. Because it is quite difficult to do a true randomized experiment in schools, such controls are key to being able to compare teachers and students.

Research on General Teacher Preparation

It is useful to consider research on general teacher preparation because this research provides some indication of how teacher preparation has been studied. In the current climate of increased accountability through NCLB, teacher preparation programs have not escaped scrutiny. With an emphasis on teacher quality and qualifications, the question of how much teacher preparation programs contribute to the learning and development of teachers has been raised. The push toward accountability in teacher preparation has several champions (Berry, 2004; U.S. Department of Education, 2003).

At the behest of the U.S. Department of Education, Wilson et al. (2001) conducted research to identify what was currently known about general teacher preparation through examining “rigorous, peer-reviewed research.” They also sought to identify gaps in the research. After examining more than 300 studies, they concluded that 57 studies met their criteria for inclusion in the report. They noted, “Overall, the research base concerning teacher preparation is relatively thin [...] good research can be done, but [...] it will take the development of more refined databases, measures, and methods, as well as complementary research designs that collect both qualitative and quantitative data” (p. i).

Most studies that make the connection between teachers’ learning and student achievement focus on subject-matter competence. For example, researchers have found that mathematics teachers who have a major in mathematics contribute to higher student mathematics scores (see, for example, Fetler, 1999; Monk, 1994). Most teacher preparation studies use teachers’ credentials as the indicator of completion of a program of study. Wilson et al. (2001) point out that this is a “weak proxy for pedagogical preparation” (p. ii). The problem with using credentials is that most studies assume that all credentials are the same (i.e., the quality, quantity, and type of pedagogical preparation is not considered, only its presence or absence). An example of this would be treating a program that requires students to take coursework to prepare them to work with students with special needs in the general classroom the same as a program that requires virtually no training in how to work with students with special needs. Thus, unless researchers specifically consider and compare the actual coursework involved in obtaining the credential, there can be wide, unexplained variations in teacher preparation that are not accounted for in the research.

Wilson et al. (2001) state, “We need more studies that relate specific parts of teachers’ preparation (subject matter, pedagogy, clinical experiences) to the effects on their teaching practice, and perhaps on student achievement” (p. iv). This dearth of teacher preparation → teacher practices → student outcomes research is obvious in examining the preparation of teachers for general education. Some studies have attempted to sort out teacher effects by looking at the same students over time as they are assigned to different teachers.

Rockoff (2004) used a random-effects meta-analysis to measure the variance of teacher fixed effects. Using elementary school teachers linked with students from two school districts in New Jersey, he examined test scores for a three-year period. He also conducted classroom observations as part of his research. He found that there were large differences in quality among teachers within schools. Such information may be useful in verifying that teacher quality is

crucial to student learning, and that some teachers are able to produce better learning outcomes than are other teachers. However, the study is unable to tell us about the relationship between teacher quality and the preparation programs that these teachers completed or about the nature or prevalence of specific practices that teachers used in their classrooms that may have contributed to student learning. Thus, the Rockoff study is an example of a study that carefully describes and confirms what most of us already suspected but provides little information that would allow us to determine how to improve teacher preparation and practices.

Research on the Teacher Preparation→ Teacher Practices→ Student Outcomes Relationship for Students With Special Needs and Their Teachers

Brownell et al. (2005) contend:

The field of special education does not have the same extensive research base on teacher quality [compared to general education], particularly as it relates to student achievement gains. Less than a handful of studies have examined linkages between dimensions of teacher quality and student achievement in education. (p. 2)

There are many studies of special education instruction, but none of them adequately connects teacher preparation→ to classroom practices→ to student learning. Some of them contribute to our understanding of pieces of this relationship, but of course, they do not help us to see the entire picture.

A Look at Several Studies

Special education teachers' ability to contribute to positive outcomes among special-needs students is a function of many factors. Carlson, Lee, and Schroll (2004) designed a study to examine special education teacher quality and identified five factors that appear to contribute to teacher quality: experience, credentials, self-efficacy, professional activities, and selected classroom practices. Using factor analysis on a sample of 1,475 special education teachers, they found that these factors contributed to an aggregate measure of teacher quality. While this is indeed an important and useful finding, it is worth pointing out that there are few studies that have attempted to define teacher quality by specific practices, a major step forward in terms of making the teacher preparation→ teacher practices→ student outcomes connection.

A study by Richards and Morse (2002) is an example of a case study that examined one preservice teacher's experience as she worked with special and general education students to develop their reading and writing skills. Although the research may be useful in understanding the development of the particular teacher and may provide insights applicable to some other teachers, it tells us little about her preparation, her practices, or the measurable outcomes of her teaching.

Another example of a study of special education teachers' preparation is a program evaluation reported by Aksamit (1990) that also used qualitative methods, in this case naturalistic inquiry methodology. The focus of the evaluation was to examine a teacher education program that merged elementary and special education into one major. Six evaluators conducted three site visits, interviewing and observing the faculty, alumni, state department of education, university personnel, and school administrators. The author does a creditable job of explaining the data collection, analysis, and coding procedures; and the findings (described as four themes) are interesting. However, we are not able to learn anything from this research that will help us understand the relationship between the program's instruction, teachers' practices, and student outcomes.

To examine the effectiveness of a program that provided rural educators with graduate-level coursework in special education, Grisham-Brown, Collins, and Baird (2000) asked teachers taking courses in special education teaching methods to examine a list of practices and to indicate which ones they had used before the course, which ones they used after the course, how many students were affected, and with how many other school staff they had shared practices and other course information. The authors were able to relate the practices used to specific teacher-preparation coursework. With 28 responses (21.7 percent), they concluded that more teachers were implementing best practices after the course and were sharing those practices with other adults. This study is a good example of how one might design research to examine the relationship between what teachers were being taught in a specific course and whether they had in fact used the practices that they had learned in their classrooms. Concerns about the limitations of self-report data and low survey response rates apply. Nonetheless, these researchers should be commended for doing something not seen often in special education research: clarifying what practices teachers should be learning in a specific course, then asking teachers to consider whether they were actually using those practices in their classrooms.

Lyon, Vaassen, and Toomey (1989) surveyed teachers' perceptions of their special education or general education teacher preparation programs. The researchers learned that although the teachers believed their programs had not provided effective instruction, they felt generally well prepared. This disparity led the researchers to conduct interviews with 30 of the teachers to clarify their seemingly contradictory perceptions. The study gives insight to how teachers regard their preparation programs after they have been teaching for some time. Unfortunately, it sheds little light on the actual connections between teacher preparation, classroom practices, and student learning.

The study that comes closest to looking at all three components of the teacher preparation → teacher practices → student outcomes relationship is that of Miller (1991). The author used a case-study approach to evaluate a project designed to facilitate the gradual integration of the special education and English teacher preparation programs in one institution. The goal of the project was to create and evaluate a model of how to integrate special education programs with general education teacher preparation programs. The author evaluated the model in several stages, and participating teachers field-tested practices they had learned in particular units. The videotaped field tests were then evaluated to determine whether the unit was taught effectively, whether the target students reached the instructional goals set for them (measured using pretests and posttests), and whether the teachers felt an increased sense of competency as a result of implementing these practices. After the evaluation of a unit was complete, the results of the evaluation were used to revise the unit to address inadequacies that had been identified. This study is especially useful since it makes two important connections that are missing in other studies: (1) what preservice teachers learned in their coursework is connected with their actual classroom practices, and (2) the connection between these practices and students' learning, measured using a pretest and posttest design, focuses on the unit being taught. In research terms, this is a major step toward looking at the impact of teacher preparation, teacher practices, and student learning.

Blackorby et al. (2005) have reported on a six-year study that provides useful information about elementary and middle-school students with disabilities through the Special Education

Elementary Longitudinal Study (SEELS). The chief advantage of this study is that it includes more than 11,000 students in a nationally representative sample. SEELS looks at the achievements of students with various disabilities. Rather than focusing solely on test scores, the outcome measures used include the following:

- **School engagement:** attending school and being actively engaged in learning activities there.
- **Academic performance:** gaining proficiency in reading, in mathematics, and in making progress in the curriculum.
- **Social adjustment:** exhibiting social skills, being socially integrated, and avoiding negative behavior.
- **Independence:** demonstrating skills that support emerging independence and assuming responsibilities at home.

Students' academic performance in reading and mathematics was assessed with measures that allow for comparisons with the general population. The usefulness of this study lies primarily in providing high-quality descriptive information on the average functioning of elementary and middle school students with a variety of disabilities. This information will be quite useful to researchers who seek to determine whether specific teacher practices are likely to, or are, making an impact on student learning. This study, however, provides no helpful information on teachers' preparation programs and the practices that teachers use in their classrooms, or on the connection between these practices and student learning.

Another study using SEELS data comes closer to answering questions about classroom practices. Wagner and Blackorby (2004) analyzed the data to determine what specific types of instruction students with disabilities were receiving in both special education language arts and general education language arts classes. For example, they found that students in general education classes were more likely to receive instruction in a whole-class group than students in special education classes (75 percent versus 49 percent). In special education classes, students were more likely to have individual instruction or to participate in small-group instruction. A particularly interesting finding from this study was that almost one third of the students taught in general education settings had teachers who reported that they were not adequately trained to teach special-needs students.

More than two decades ago, a study by Englert (1984) examined teachers' direct instruction practices in special education settings. The Direct Instructional Observation System (DIOS) was used to code teachers' instructional practices. Students' pretest and posttest scores were subsequently used to divide teachers into two effectiveness groups. Statistical methods were used to compare the teaching skills and strategies of the two groups of teachers. Englert found significant differences. She also found that three variables in particular made a significant contribution to the differences in outcomes. More effective teachers had higher occurrences of lesson objectives, concept examples, and error drills. This study is a particularly good example of mixed-methods research, where the qualitative component (used to collect data on teacher behaviors in the classroom) and the quantitative component (used to collect pretest and posttest data from students and to analyze differences in two groups of teachers) are both critically

important to the study. The design of the study prefigures the current emphasis on determining teacher effectiveness by outputs (student achievement) rather than by inputs (such as teachers' credentials and education). Although this study does make the link between teacher practices and student learning, it does not connect teacher practices to their teacher-preparation experiences.

Another older study that merits discussion is Brophy (1985), which provides a useful example of a research review on a specific topic—teacher behaviors as they relate to student achievement. For a chapter in the *Handbook of Research on Teaching*, Brophy examined studies to look for factors that were consistently found to contribute to student achievement, regardless of the size of the correlation. Certain factors such as the quantity and pacing of instruction appeared to affect student achievement. This is a useful study in that it accomplishes an important purpose: informing researchers and practitioners about the types of teacher behaviors that consistently appear to impact student achievement across all types of studies in a variety of settings. However, because of the nature of such aggregation, much of the specificity is lost in the Brophy-reviewed studies. Thus, while we learn that *active teaching* rather than *individual work* appears to be consistently related to student achievement, what we really want to know is *which* specific practices under the general heading of active teaching is the teacher using? To be able to trace the practices back to the teacher preparation program, we would need to know more than simply that the teachers were taught to engage in active teaching.

The difficulties in conducting research that connects teacher preparation with actual classroom practices lie primarily in evaluating teachers' practices in the classroom. Some studies on teachers who completed coursework while already teaching have attempted to use self-report surveys as a measure of change-in-practice.

There are a number of difficulties in conducting such a study and in interpreting the results. First, mail-in response rates are typically low. In this study, no follow-up was done with teachers who failed to respond; but even with intensive follow-up, return rates for mailed surveys seldom exceed 30 percent to 35 percent. In addition, in this study there was no way of knowing whether those who returned the surveys were representative of all the participating teachers. And because the data is self-reported, there is no way to verify whether teachers are accurately characterizing the changes they made in their teaching practices.

Blanton, Blanton, and Cross (1994) conducted a well-designed mixed-method study of how general and special education teachers make instructional decisions and noted that their findings have implications for teacher preparation programs. This is a sort of backward mapping (i.e., looking at the practices and then connecting them to what teachers *should be* learning in their teacher preparation programs). The teachers in the study were asked to watch a video of a reading lesson (developed by the researchers) and then to respond to 45 items (in an instrument developed by the researchers), rating them on a Likert scale from 1 (not important) to 4 (very important). They also completed an open-ended probe as they watched the video, describing the target student's strengths and needs as well as strategies that might help the student do well in a general classroom setting. The teachers also were asked to tell why the strategies they recommended were important and to describe their role in implementing the strategies. The subjects included 20 special education resource teachers and 20 general education teachers,

matched by experience and other factors. The researchers then analyzed the teachers' responses and looked for significant differences between the two groups of teachers.

The important findings from Blanton et al. (1994) were: (1) the special education teachers were more likely to emphasize curriculum and instructional variables in their assessment of the student's performance, while the general education teachers were more likely to emphasize social and behavioral variables and classroom management; (2) special education teachers possessed greater knowledge when it came to identifying possible problems and suggesting responses to them; and (3) neither of the two groups of teachers indicated *conferring with each other* as a possible strategy. These findings appear to point out issues that should probably be addressed in teacher preparation programs.

While this study is designed, conducted, and analyzed well and it provides some interesting information about the differences in how special and general education teachers react to struggling students, it does not specifically connect this information to the practices these teachers learned in their teacher preparation program; although the researchers do suggest that this is useful information for considering changes in that preparation curriculum. However, this may involve tracing the *absence* of practices back to the teacher preparation program—a greater challenge than tracing existing practices back to the program. The research also does not make any connection with student outcomes (i.e., if the teachers' practices were changed, would students benefit from those changes?). Incorporating that type of question might be useful in driving future research.

Recommendations

Because of the dearth of research on the teacher preparation→ teacher practices→ student outcomes relationship for students with special needs and their teachers, it is not possible to list best practices or to draw conclusions that will help teachers in their day-to-day instruction. However, looking at gaps in research can provide useful information that can help shape policy and impact teachers and students in positive ways. It is possible to see patterns in the gaps. These patterns can provide useful information about how subsequent research can be approached. By examining existing research and noting where additional research could extend the findings to include all the links in the teacher preparation→ teacher practices→ student outcomes connection, it is possible to see future directions that researchers might take.

1. Encourage and Support Research

Encourage and support research into the teacher preparation→ teacher practices→ student outcomes relationship for students with special needs and their teachers, particularly research that ties specific teacher practices to specific coursework and ties specific teacher practices to student learning outcomes. Support in this instance means prioritizing funding to ensure that researchers have sufficient means to undertake complex research that may require a mixed-methods design. Such research involves more time and more person-hours to complete. The ideal research design would involve researchers working directly with states to examine longitudinal data linking teacher characteristics and qualifications to student outcomes.

The first and most obvious need in research into the teacher preparation→ teacher practices→ student outcomes relationship for students with special needs and their teachers is for more (and more rigorous) research to be conducted that takes into account all three components. Studies have been conducted that have shown how parts of this research might be carried out, as has been shown in this report. Wilson et al. (2001) state, “We need more studies that relate specific parts of teachers’ preparation (subject matter, pedagogy, clinical experiences) to the effects on their teaching practice, and perhaps on student achievement. Studies that compare the relative importance of specific parts of teacher preparation could be useful to those designing and revising teacher education programs” (p. iv). As these authors suggest, ensuring that teacher preparation programs are actually producing effective teachers is only half the story. We also need specific data that will allow teacher preparation programs to retool their course offerings and curricula to ensure that what teacher candidates are learning will eventually make meaningful contributions to student outcomes.

2. Extend the Research to Include Induction and Early Professional Development

Because teacher learning continues to take place steadily even after teachers have completed their preparation programs, the focus of research on new teacher preparation should be extended to include induction and early professional development.

Feiman-Nemser (2001) has argued that we need to think beyond teacher preparation programs and start conceptualizing teacher learning as a continuum. She suggests a “framework for

thinking about a curriculum for teacher learning over time” (p. 1013). From this perspective, teacher preparation programs would be only the first part of a continuum that includes new teacher induction and early-career professional development. This makes sense from a research perspective as well because new teacher learning certainly does not happen only in teacher preparation programs. Teacher change is a complex and ongoing process, influenced by teachers’ beliefs and justifications about the need to adopt specific practices (Richardson, 1990).

Thus, when examining teacher practices to relate them to teacher preparation and student outcomes, the ability to clearly determine the specific influences on those teacher practices is important. For example, it is important to know if the practice was learned in a teacher preparation program, from fellow teachers, from a mentor teacher, in an induction program, or through professional development activities.

3. Consider Curriculum-Based Measures of Achievement for Students With Special Needs

Deno, Fuchs, Marston, and Shen (2001) have suggested measuring special education student achievement in the classroom with curriculum-based measurements (CBM) rather than standardized achievement tests. CBMs are commonly used to measure reading ability, and there is a high correlation between CBMs measuring oral fluency and comprehension (Deno, 2003). Because CBMs are a relatively quick, easy, and inexpensive way to measure student progress and because they are highly correlated with other indicators of student learning, they hold great promise as an outcome measure that can be tied directly to specific teacher practices.

Increasing the use of CBMs will resolve several problems with doing research on the teacher preparation → teacher practices → student outcomes relationship for students with special needs and their teachers. First, it would allow measurement of just the part of students’ learning that can be tied directly to a particular teacher. As previously noted, there are typically several teachers involved in the instruction of students with special needs. With CBMs, it is possible to focus on small, specific learning objectives in one subject area, thus making it easier to tie student learning to specific teachers’ specific practices. This is only the beginning, of course. It is important to show where in a teacher’s own learning continuum he or she learned the particular practice, and with what consistency it has been practiced. Furthermore, aggregating results over many teachers and students would be necessary to be able to reach defensible conclusions about how the particular practice affects student learning.

4. Include Special Education Instruction in all Teacher Preparation Programs, and Encourage Collaboration Between General and Special Education Teachers

Focus preparation programs to include instruction in special education for general education teachers as well as special education teachers and provide explicit instruction for collaborations among special and general education teachers.

This is particularly crucial in the early identification of struggling students who may need immediate interventions. Response to intervention or other approaches that seek out and attempt

to rectify learning challenges for individual students are best conducted collaboratively rather than by a special or general education teacher acting independently. Furthermore, collaborative efforts by general and special education instructors will ensure that the changing realities for special-needs students (such as inclusion in standardized testing and grade-level instruction) will be more effectively managed.

5. Design Research That Measures Collective Contributions to Student Outcomes

In designing research into the teacher preparation → teacher practices → student outcomes relationship, consider design features that take into account the multiple teacher contributions toward student learning.

Because most students with special needs work with several teachers throughout the day, it is difficult, perhaps impossible, to sort out the individual contributions made by specific teachers. However, it should be possible to conduct research that measures a collective teacher contribution that ties all a particular student's teachers to that student's outcomes.

6. Develop a Comprehensive Longitudinal Teacher Database

Develop a state-level longitudinal database that includes background information about teachers' education such as where (and when) they attended college and where (and when) they completed coursework for certification. Include information on alternative certification programs as well.

Most states are working toward developing longitudinal databases to help them meet requirements of NCLB, particularly the highly qualified teacher component. States must be able to show that teachers are highly qualified in the subjects they are teaching. This means that states must be able to link teachers to actual classroom assignments. As states are building the infrastructure of documents to verify that their teachers are highly qualified, they may want to add additional teacher information that will assist them in doing research on other aspects of teacher quality and qualifications. Additionally, if states (or researchers working collaboratively with states) want to attempt to link specific teacher qualifications to student achievement, then a longitudinal database must be created that also links teachers to students. The reason that the data must be maintained in a longitudinal database is that information about students' prior achievement is often used as a control in growth models, and several years of data linking student test scores to teachers are necessary for valid and useful value-added modeling.

Conclusion

There is little doubt that a more thorough study of the teacher preparation → teacher practices → student outcomes relationship for students with special needs and their teachers would be beneficial for two reasons. First, we need to know much more about what teachers are learning in preparation programs that they are then implementing in their classrooms to further student learning. Second, such research would provide empirical evidence on the specific practices or types of practices that contribute most to student success. Although some research exists in this area, there is little research that provides strong evidence that one specific practice is better than another in terms of student outcomes. With hard evidence on what teachers learn in their preparation programs, what they subsequently do in their classrooms, and what their students learn as a result, we will have the information we need to work toward creating better teachers, better teaching, and better outcomes.

References

- Achinstein, B., & Barrett, A. (2004). (Re)framing classroom contexts: How new teachers and mentors view diverse learners and challenges of practice. *Teachers College Record*, 106(4), 716–746.
- Aksamit, D. L. (1990). Naturalistic inquiry applied to the evaluation of a teacher education program. *Teaching and Teacher Education*, 6(3), 215–226.
- Allen, M. B. (2003). *Eight questions on teacher preparation: What does the research say? A summary of the findings* (ERIC Document No. ED479051). Denver, CO: Education Commission of the States. Retrieved October 27, 2006, from http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/22/d0/0b.pdf
- Allen, M., & Palaich, R. (2000). In pursuit of quality teaching: Five key strategies for policymakers. Denver, CO: Education Commission of the States.
- American Federation of Teachers. (2000). *Building a profession: Strengthening teacher preparation and induction*. Washington, DC: Author. Retrieved October 27, 2006, from <http://65.110.81.56/pubs-reports/downloads/teachers/k16report.pdf>
- Belden Russonello & Stewart. (2003). *Expectations, achievements and challenges teaching special education students*. Washington, DC: Education Week. Retrieved October 27, 2006, from http://counts.edweek.org/sreports/qc04/tables/Teacher_poll_report.pdf
- Berry, B. (2004). *Making good on what matters most: A review of teaching at risk: A call to action* (Analysis). Chapel Hill, NC: Southeast Center for Teaching Quality
- Blackorby, J., Wagner, M., Cameto, R., Davies, E., Levine, P., Newman, L. et al. (2005, October). *Engagement, academics, social adjustment, and independence: The achievements of elementary and middle school students with disabilities*. Menlo Park, CA: SRI International. Retrieved October 27, 2006, from http://www.seels.net/designdocs/engagement/All_SEELS_outcomes_10-04-05.pdf
- Blanton, L. P., Blanton, W. E., & Cross, L. S. (1994). An exploratory study of how general and special education teachers think and make instructional decisions about students with special needs. *Teacher Education and Special Education*, 17(1), 62–74.
- Blanton, L. P., Griffin, C. C., Winn, J. A., & Pugach, M. C. (Eds.). (1997). *Teacher education in transition*. Denver, CO: Love Publishing Company.
- Brewster, C., & Klump, J. (2005). *Leadership practices of successful principals*. Portland, OR: Northwest Regional Educational Laboratory. Retrieved October 27, 2006, from <http://www.nwrel.org/request/2005nov/leadership.pdf>

- Brophy, J., & Good, T. (1985). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.
- Brownell, M. T., Bishop, A. G., Gersten, R., Klingner, J. K., Dimino, J., Haager, D., et al. (2005). *Defining and assessing the quality of beginning special education teachers: First steps, conclusions drawn, and challenges encountered*. (2005, August). Paper presented at the Inclusive and Supportive Education Congress International Special Education Conference. Glasgow, Scotland. Retrieved October 27, 2006, from http://www.isec2005.org.uk/isec/abstracts/papers_b/brownell_m_2.shtml
- Bryk, A. S., & Raudenbush, S. W. (1988, November). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. *American Journal of Education*, 97(1), 65–108.
- Burstein, L. (1980) The analysis of multilevel data in educational research and evaluation. In D. C. Berliner (Ed.), *Review of research in education*, vol. 8, (pp. 158–233). Washington, DC: American Educational Research Association.
- Carlson, E., Lee, H., & Schroll, K. (2004). Identifying attributes of high quality special education teachers. *Teacher Education and Special Education*, 27(4), 350–359.
- Carnine, D. (1999). Perspective: Campaigns for moving research into practice. *Remedial and Special Education*, 20(1), 2–6.
- Cochran-Smith, M., & Zeichner, K. M. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. Mahwah, NJ: Erlbaum.
- Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record*, 102(2), 294–343.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cook, T. D. (2002). Randomized experiments in education: Why are they so rare? *Educational Evaluation and Policy Analysis*, 24(3), 175–200.
- Cooper, H., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.
- Corcoran, T. B., Walker, L. J., & White, J. L. (1988). *Working in urban schools*. Washington, DC: Institute for Educational Leadership.

- Cronbach, L. J., Linn, R. L., Brennan, R. L., & Haertel, E. H. (1997). Generalizability analysis for performance assessments of student achievement or school effectiveness. *Educational and Psychological Measurement, 57*(3), 373.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1). Retrieved October 27, 2006, from <http://epaa.asu.edu/epaa/v8n1/>
- Data Quality Campaign. (2005). *Creating a longitudinal data system: Using data to improve student achievement*. Washington, DC: Author. Retrieved October 27, 2006, from http://www.dataqualitycampaign.org/files/Publications-Creating_Longitudinal_Data_System.pdf
- Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*(3), 184–192.
- Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*(4), 507–524.
- Doran, H. C., & Izumi, L. T. (2004). *Putting education to the test: A value-added model for California*. San Francisco: Pacific Research Institute. Retrieved October 27, 2006, from http://www.pacificresearch.org/pub/sab/educat/2004/Value_Added_factsheet.pdf
- Dwyer, C. A., Millett, C. M., & Payne, D. G. (2006). *A culture of evidence: Postsecondary assessment and learning outcomes*. Princeton, NJ: ETS. Retrieved October 27, 2006, from http://www.ets.org/Media/Resources_For/Policy_Makers/pdf/cultureofevidence.pdf
- Eisenhart, M., & Towne, L. (2003). Contestation and change in national policy on “scientifically based” education research. *Educational Researcher, 32*(7), 31–38. Retrieved October 27, 2006, from http://www.aera.net/uploadedFiles/Journals_and_Publications/Journals/Educational_Researcher/3207/3207_Eisenhart.pdf
- Englert, C. S. (1984). Effective direct instruction practices in special education settings. *Remedial and Special Education, 5*(2), 38–47.
- Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013–1055.
- Feistritzer, C. E. (1999). *The making of a teacher: A report on teacher preparation in the U.S.* Washington, DC: National Center for Education Information. Retrieved October 27, 2006, from <http://www.ncei.com/MakingTeacher-rpt.htm>
- Ferguson, P., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education, 44*(1), 55–63.

- Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives*, 7(9). Retrieved October 27, 2006, from <http://epaa.asu.edu/epaa/v7n9.html>
- Gersten, R., Baker, S., & Pugach, M. (with Scanlon, D., & Chard, D.). (2001). Contemporary research on special education teaching. In V. Richardson (Ed.), *Handbook of Research on Teaching* (4th ed., pp. 695–722). Washington, DC: American Educational Research Association.
- Goe, L. D., Thompson, M. R., Paek, P. L., & Ponte, E. (2005, April). Context and implementation issues: A study of a statewide teacher induction program (CFASST). Paper presented at the annual meeting of the American Educational Research Association, Montreal. Retrieved October 27, 2006, from http://www.ets.org/Media/Resources_For/K-12_Education/pdf/AERA_2005_Thompson.pdf
- Goertz, M. E. (2000, March). *The state role in education accountability*. Paper presented at the EdSource Conference, San Ramon, CA.
- Goertz, M. E., & Duffy, M. C. (with K. Carlson Le Floch). (2000). *Assessment and accountability systems in the 50 states: 1999–2000*. Philadelphia: Consortium for Policy Research in Education. Retrieved October 27, 2006, from <http://www.cpre.org/Publications/rr46.pdf>
- Goldhaber, D. D., & Brewer, D. J. (Eds.). (1999). Teacher licensing and student achievement. In M. Kanstoroom & C. E. Finn (Eds.), *Better teachers, better schools* (pp. 83–102). Washington, DC: Thomas B. Fordham Foundation. Retrieved October 27, 2006, from <http://www.msu.edu/user/mkennedy/TQQT/PDFs/GoldBrew%2099.pdf>
- Gong, B. (2004, November). *Models for using student growth measures in school accountability*. Paper presented at the 2004 CCSSO Brain Trust, Washington, DC. Retrieved October 27, 2006, from <http://www.nciea.org/publications/GongGrowthModels111504.pdf>
- Griffin, C. C., & Pugach, M. C. (1997). Framing the progress of collaborative teacher education. In L. P. Blanton, C. C. Griffin, J. A. Winn & M. C. Pugach (Eds.), *Teacher education in transition: Collaborative programs to prepare general and special educators* (pp. 249–270). Denver: Love Publishing.
- Grisham-Brown, J., Collins, B. C., & Baird, C. M. (2000). *Training rural educators in Kentucky through distance learning: Impact with follow-up data*. Paper presented at Capitalizing on Leadership in Rural Special Education: Making a Difference for Children and Families Conference, Alexandria, VA. Retrieved October 27, 2006, from http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/10/ce/48.pdf
- Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Mahwah, NJ: Erlbaum.

- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). *Teachers, schools, and academic achievement* (Working Paper No. 6691). Cambridge, MA: National Bureau of Economic Research.
- Hardman, M. L., McDonnell, J., & Welch, M. (1998). *Special education in an era of school reform: Preparing special education teachers*. Washington, DC: Federal Resource Center for Special Education. Retrieved October 27, 2006, from <http://www.dssc.org/frc/pubs/hardman.pdf>
- Hattie, J. A. (1992). Towards a model of schooling: A synthesis of meta-analysis. *Australian Journal of Education*, 36, 5–13.
- Heubert, J. P., & Hauser, R. M. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- Johnson, B. (2001). Toward a new classification of nonexperimental quantitative research. *Educational Researcher*, 30(2), 3–13. Retrieved October 27, 2006, from http://www.aera.net/uploadedFiles/Journals_and_Publications/Journals/Educational_Researcher/3002/AERA3002_Johnson.pdf
- Kavale, K. A. (2005). Effective intervention for students with specific learning disabilities: The nature of special education. *Learning Disabilities: A Multidisciplinary Journal*, 13(4), 127–138.
- Kirby, S. N., McCaffrey, D. F., Lockwood, J. R., McCombs, J. S., Naftel, S., & Barney, H. (2002). Using state school accountability data to evaluate federal programs: A long uphill road. *Peabody Journal of Education*, 77(4), 122–145.
- Koretz, D. (2000). *Limitations in the use of achievement tests as measures of educators' productivity*. Santa Monica, CA: RAND Education Center for Research on Evaluation, Standards, and Student Testing. Retrieved October 27, 2006, from <http://www.irp.wisc.edu/newsevents/other/symposia/koretz.pdf>
- Kupermintz, H. (2003). Teacher effects and teacher effectiveness: a validity investigation of the Tennessee value added assessment system. *Educational Evaluation and Policy Analysis*, 25(3), 287–298.
- Ladd, H. F. (2001). School-based education accountability systems: The promise and the pitfalls. *National Tax Journal*, 54(2), 385–400.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4–16.
- Lyon, G. R., Vaassen, M., & Toomey, F. (1989). Teacher perceptions of their undergraduate and graduate preparation. *Teacher Education and Special Education*, 12(4), 164–169.

- Maheady, L. (1997). Preparing teachers for instructing multiple ability groups. *Teacher Education and Special Education*, 20(4), 322–339.
- McAllister, G. (2000). *Does context matter? Urban teachers' beliefs about culture and the relationship of those beliefs to their practice*. Unpublished manuscript, Flagstaff, AZ.
- McCaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 32(5), 493–517.
- McDonnell, L. M., McLaughlin, M. J., & Morison, P. (1997). *Educating one and all: Students with disabilities and standards-based reform*. Washington, DC: National Academy Press.
- McLaughlin, M. W., Talbert, J. E., & Bascia, N. (1990). *The contexts of teaching in secondary schools: Teachers' realities*. New York: Teachers College Press Columbia University.
- Mendro, R. L., Jordan, H. R., Gomez, E., Anderson, M. C., & Bembry, K. L. (1998, April). *An application of multiple linear regression in determining longitudinal teacher effectiveness*. Paper presented at the 1998 annual meeting of the American Educational Research Association, San Diego, CA.
- Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, 79(4), 33–53.
- Miller, D. E. (1991). Merging regular and special education teacher preparation programs: The integrated special education-English project (ISEP). *Teaching and Teacher Education*, 7(1), 19–23.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13(2), 125–145.
- Mullens, J. E., Murnane, R. J., & Willett, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review*, 40(2), 139–157.
- National Association of State Directors of Special Education (NASDSE), & Council of Administrators of Special Education (CASE). (2006). *Response to intervention: NASDSE and CASE white paper on RtI* (white paper). Alexandria, VA: Author. Retrieved October 27, 2006, from <http://www.nasdse.org/documents/RTIAnAdministratorsPerspective1-06.pdf>
- National Commission on Excellence in Education. (1983, April). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office. Retrieved October 27, 2006, from <http://www.ed.gov/pubs/NatAtRisk/index.html>

- Noell, G. H. (2006). *Value added assessment of teacher preparation* (Annual Report). Baton Rouge, LA: Louisiana Board of Regents. Retrieved October 27, 2006, from <http://asa.regents.state.la.us/TE/2005-technical-report.pdf>
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257.
- Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*(2), 137.
- Pugach, M. C. (2005). Research on preparing teachers to work with students with disabilities. In M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 549–590). Mahwah, N.J.: Erlbaum.
- Pugach, M. C., & Seidl, B. L. (1995). From exclusion to inclusion in urban schools. *Education and Urban Society, 27*(4), 379–395.
- Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher, 34*(5), 25–31.
- Richards, J. C., & Morse, T. E. (2002, June). One preservice teacher's experiences teaching literacy to regular and special education students. *Reading Online, 5*, 1–16.
- Richardson, V. (1990). Significant and worthwhile change in teaching practice. *Educational Researcher, 19*(7), 10–18.
- Riehl, C. (2006). Feeling better: A comparison of medical research and education research. *Educational Researcher, 35*(5), 24–29. Retrieved October 27, 2006, from http://www.aera.net/uploadedFiles/Publications/Journals/Educational_Researcher/3505/3790-03_Riehl-RN&C.pdf
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review, 94*(2), 247–252.
- 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*(8), 1525–1567.
- Sanders, S. L., Skonie-Hardin, S. D., Phelps, W. H., & Minnis, T. L. (1994, November). *The effects of teacher educational attainment on student educational attainment in four regions of Virginia: Implications for administrators*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Nashville, TN. Retrieved October 27, 2006, from http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/22/3e/06.pdf

- Sanders, W. L., & Horn, S. P. (1995). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Education Policy Analysis Archives*, 3(6). Retrieved October 27, 2006, from <http://epaa.asu.edu/epaa/v3n6.html>
- Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247–256.
- Sanders, W. L., & Rivers, J. (2006, April) The value of evidence in reforming teacher education. *Teachers for a New Era Quarterly Newsletter*, 2(3), 1–2. Retrieved October 27, 2006, from <http://www.teachersforanewera.org/newsletters/newsletters/TNE%20Newsletter%20V2-N3.doc>
- Sanders, W. L., Saxton, A. M., & Horn, S. P. (1997). The Tennessee value-added assessment system: A quantitative, outcomes-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 137–162). Thousand Oaks, CA: Corwin Press.
- Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11(1), 57–67.
- Schochet, P. Z. (2005). *Statistical power for random assignment evaluations of education programs*. Princeton, NJ: Mathematica Policy Research. Retrieved October 27, 2006, from <http://www.mathematica-mpr.com/publications/PDFs/statisticalpower.pdf>
- Schofield, J. W., & Sagar, H. A. (1979). The social context of learning in an interracial school. In R. C. Rist (Ed.), *Desegregated schools: Appraisals of an American experiment* (pp. 155–199). New York: Academic Press.
- Scribner, J. P. (1999). Professional development: Untangling the influence of work context on teacher learning. *Educational Administration Quarterly*, 35(2), 238–266.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference* (1st ed.). Boston: Houghton Mifflin.
- Shann, M. H. (1998). Professional commitment and satisfaction among teachers in urban middle schools. *Journal of Educational Research*, 92(2), 67–73.
- Shulman, L. S. (2005). Teacher education does not exist. *Stanford Educator*, Fall '05, 7. Retrieved October 27, 2006, from <http://ed.stanford.edu/suse/news-bureau/educator/fall2005/EducatorFall05.pdf>
- Slavin, R. E. (2004). Education research can and must address “what works” questions. *Educational Researcher*, 33(1), 27–28.

- The Teaching Commission. (2004). *Teaching at risk: A call to action*. New York: Author. Retrieved October 27, 2006, from <http://ftp.ets.org/pub/corp/ttcreport.pdf>
- Thompson, M., Paek, P., Goe, L., & Ponte, E. (2004). *Study of the impact of the California formative assessment and support system for teachers: Report 3, relationship of BTSA/CFASST engagement and student achievement*. Princeton: ETS. Retrieved October 27, 2006, from <http://ftp.ets.org/pub/res/researcher/RR-04-32.pdf>
- U.S. Department of Education. (2003). *Meeting the highly qualified teachers challenge: The secretary's second annual report on teacher quality*. Washington, DC: Author. Retrieved October 27, 2006, from <http://www.ed.gov/about/reports/annual/teachprep/2003title-ii-report.pdf>
- Vandevort, L. G., Amrein-Beardsley, A., & Berliner, D. C. (2004). National board certified teachers and their students' achievement. *Education Policy Analysis Archives*, 12(46). Retrieved October 27, 2006, from <http://epaa.asu.edu/epaa/v12n46/>
- Wagner, M., & Blackorby, J. (2004). *Overview of findings from wave 1 of the special education elementary longitudinal study (SEELS)*. Menlo Park, CA: SRI International. Retrieved October 27, 2006, from http://www.seels.net/designdocs/seels_wave1_9-23-04.pdf
- Webster, W. J., & Mendro, R. L. (1997). The Dallas value-added accountability system. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 81–99). Thousand Oaks, CA: Corwin Press.
- What Works Clearinghouse. (2003). *Introduction to the What Works Clearinghouse: WWC evidence report process and the role of scientific standards*. Rockville, MD: Author.
- Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations*. Seattle, WA: Center for the Study of Teaching and Policy. Retrieved October 27, 2006, from <http://depts.washington.edu/ctpmail/PDFs/TeacherPrep-WFFM-02-2001.pdf>
- Wineburg, M. S. (2006). Evidence in teacher preparation: Establishing a framework for accountability. *Journal of Teacher Education*, 57(1), 51–64.
- Ysseldyke, J., Dennison, A., & Nelson, R. (2004). *Large-scale assessment and accountability systems: Positive consequences for students with disabilities* (NCEO Synthesis Report 51). Minneapolis, MN: National Center on Educational Outcomes. Retrieved October 27, 2006, from <http://www.education.umn.edu/nceo/OnlinePubs/Synthesis51.html>