Educational Evaluation and Policy Analysis
http://eepa.aera.net

Inquiry-Oriented Instruction in Science: Who Teaches That Way?

Thomas M. Smith, Laura M. Desimone, Timothy L. Zeidner, Alfred C. Dunn, Monica Bhatt and Nataliya L. Rumyantseva

EDUCATIONAL EVALUATION AND POLICY ANALYSIS 2007 29: 169
DOI: 10.3102/0162373707306025

Educational Evaluation and Policy Analysis
September 2007, Vol. 29, No. 3, pp. 169–199
DOI: 10.3102/0162373707306025

Inquiry-Oriented Instruction in Science: Who Teaches That Way?

Thomas M. Smith, Laura M. Desimone, Timothy L. Zeidner, Alfred C. Dunn, Monica Bhatt, and Nataliya L. Rumyantseva
Vanderbilt University

The expansion of the No Child Left Behind Act to include science standards and assessments is likely to refocus states' attention on science teaching and learning. Requiring teachers to have subject majors and greater funding of professional development are two key policy levers for improving instruction in science. There has been relatively little work examining the characteristics of teachers who are most likely to initiate inquiry-oriented instruction in science classrooms. Using a nationally representative sample of the teachers of eighth grade science students, the authors found relatively strong associations between reform-oriented practice and the majors and degrees that teachers earned as part of their formal schooling, as well as their current levels of participation in content-oriented professional development activities.

Keywords: *science teaching, professional development, teacher quality*

Although most states and districts have focused their recent reform efforts on reading and mathematics, the extension of the No Child Left Behind Act of 2001 (NCLB) to include assessments of science achievement in 2007–2008 is likely to pull this subject back into the policy debate. Policy makers' concern about U.S. students' performance in science has fluctuated over the past half century, often driven by comparatively poor performance on international assessments and concerns over challenges to U.S. competitiveness in technological fields (Marx & Harris, 2006).

There has been concern among those in the science community that this periodic focus on improving students' performance in science has waned since NCLB was enacted in 2001, as state and local officials focused on the federal law's demand that they improve annual test scores in reading and mathematics in Grades 3–8 (Cavanagh, 2005; Marx & Harris, 2006). Recently, science education has come back under the microscope, as NCLB required states to have science content standards in place for the 2005–2006 school year and to start testing students yearly in science at least once in each of the 3–5, 6–9, and 10–12 grade spans in 2007–2008. Although it will be up to individual states to decide if students' performance on these assessments will be a factor in determining whether a district or school is meeting adequate yearly progress, the renewed focus on the implementation of science standards and aligned tests makes it ever more critical to understand the contexts under which science is being taught in a manner

This article is based on work supported by the National Science Foundation under Grant 0231884. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

consistent with the reform principles delineated by the science community.

Much of the push within the science community to improve science education in the United States over the past decade has involved developing new standards for what students should know and be able to do in science. The standards were not created in a vacuum; instead, they are broadly consistent with the objectives of a number of curricular projects undertaken in the 1960s and 1970s, including the Biological Sciences Curriculum Study programs in biology, the Physical Sciences Study Committee materials in physics, and the Science Curriculum Improvement Study and Elementary Science Study units for elementary school science, which incorporated approaches to teaching and learning that would fit under what science standards would call inquiry (National Research Council, 2000). The National Science Education Standards (NSES) issued by the National Research Council (1996) describe standards as "criteria to judge progress toward a national vision of learning and teaching science in a system that promotes excellence." These standards for science call for a different kind of instruction than what currently occurs in many U.S. classrooms (National Research Council, 1999). The kinds of "inquiry-oriented" instruction described in the standards treat learning as an active process, involving making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires the "identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations" (National Research Council, 1996, p. 23). Inquiry is at the center of this kind of instruction, in contrast to what has historically occurred in U.S. classrooms, in which the emphasis has been on students' ability to record information presented by teachers and to memorize scientific facts and formulas (Anderson, 2002; Cohen & Spillane, 1993; Von Secker & Lissitz, 1999). Anderson and Helms (2001) distinguished between types of inquiry in three contexts: (a) inquiry as a descriptor of scientific research, (b) inquiry as a type of teaching, and (c) inquiry as a mode of student learning. Although we recognize the importance of each of these in the NSES, this article focuses on the second, inquiry as a type of teaching, and we use the term *inquiry-oriented instruction* to describe what kind of teaching is being advocated, acknowledging that neither the field nor even the NSES share a common operational definition for what inquiry-oriented instruction is.

Although there is a limited research base suggesting that inquiry-oriented instruction is associated with improved student performance in science (described below), there has been relatively little work examining the characteristics of teachers who are most likely to initiate this kind of instruction in science classrooms. NCLB has begun to mandate the required characteristics of teachers, including (a) having a bachelor's degree; (b) being fully certified or licensed, including certification obtained through "alternative routes" (No Child Left Behind Act, 2001, sect. 9191, part 23A[i]); and (c) demonstrating content knowledge in the subjects they teach. Veteran teachers can meet the content knowledge requirement by passing state-determined tests or having college majors or through some other process defined by the states, including a combination of experience, college course work, professional development, and other state-determined measures. Professional development is a major focus of current educational reform initiatives (Corcoran, 1995; Corcoran, Shields, & Zucker, 1998; Sykes, 1996) and is foreseen as one of the primary ways that veteran teachers can attain "highly qualified" status. Within the next few years, NCLB may have a major impact on how science is taught by (a) requiring aligned state-level standards and assessments in science from elementary to high school, and (b) allowing states to include science assessment results in decisions about whether schools and districts are meeting adequate yearly. Similar accountability requirements have been associated with increases in basic rather than reform-oriented learning in math (Desimone, Smith, Hayes, & Frisvold, 2005), and it remains to be seen how such provisions will affect science learning.

NCLB will also require science teachers to be certified in science and demonstrate proficiency in their subject matter.

In this article we examine the likely impact of these latter "quality mandates" on increasing teachers' use of methods that support inquiry, as defined below, in classrooms. We do this by examining the relationship between teachers' experience levels, certification status, whether they have a degree in science, and their participation rates in science-related professional development activities on (a) the amount of time they spend on procedural activities, (b) the amount of time they spend on reporting and writing activities, (c) the amount of time they spend on hands-on activities, and (d) the amount of emphasis they give to conceptual objectives. Although the activities and objectives reported by teachers in this study are not a one-to-one match to the NSES, there is sufficient consistency between them to make it worthwhile to examine the characteristics of teachers, as well as their schools and students, that are associated with teachers placing greater emphasis on these reform-oriented instructional practices. These analyses help us assess whether policy levers currently being manipulated by states and districts to change teacher quality are at least associated with the kinds of instruction going on in science classrooms.

Prior Research in Inquiry-Oriented Teaching

According to the NSES,

> scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world. (National Research Council, 1996, p. 23)

Although inquiry-oriented approaches to education have been advocated since Dewey early in the last century (Dewey, 1910, as cited in National Research Council, 2000), what constitutes inquiry in science instruction varies widely across the literature (Anderson, 2002). For example, prior meta-analyses of the effectiveness of inquiry teaching have focused on materials (Shymansky, Kyle, & Alport, 1983); "inquiry-discovery" teaching, described as "more student-centered and less step-by-step teacher learning" (Wise & Okey, 1983); a classification of teaching as either inductive or deductive in approach (Lott, 1983); as well as what has more recently been called "project-based" science instruction (e.g., Blumfeld, Krajcik, Marx, & Soloway, 1994) and Science Education Through Portfolio Instruction & Assessment, which includes organizing lessons around solving a problem that engages students, has a clear science content focus, and emphasizes scientific reasoning ("minds-on" inquiry; Duschl & Gitomer, 1997; Gitomer & Duschl, 1995).

Although not focusing on a single operational definition of inquiry, the National Research Council (2000) lists the following eight features as essential to classroom inquiry at the middle school level:

- Identify questions that can be answered through scientific investigations.
- Design and conduct a scientific investigation.
- Use appropriate tools and techniques to gather, analyze, and interpret data.
- Develop descriptions, explanations, predictions, and models using evidence.
- Think critically and logically to make the relationships between evidence and explanations.
- Recognize and analyze alternative explanations and predictions.
- Communicate scientific procedures and explanations.
- Use mathematics in all aspects of scientific inquiry.

More recent changes in the expectations for what students should know in science can also be seen from the modifications that were made to the science frameworks for the National Assessment of Educational Progress (NAEP). Although the 1996–2005 NAEP framework organized the dimension of knowing and doing into conceptual understanding, science investigation, and practical investigation, the 2009 framework highlights four science practices to be assessed: (a) identifying science principles, (b) using science principles, (c) using scientific inquiry, and (d) and using technological design (U.S. Department of Education, 2007b). Wright and Wright (1998) pointed out the wide gap between science education as it was taught at the time and as it was described in the NSES, suggesting that it would take considerable effort for teachers and students to enact the vision of inquiry put forth in the guidelines in their classrooms.

As state science frameworks and curriculum frameworks continue to be influenced by the emphasis on inquiry in the NSES (Hollweg & Hill, 2003) and as tests based on these frameworks are increasingly likely to come on-line as part of NCLB, it is important to understand both the evidence in the literature supporting inquiry-oriented instruction as well as the contexts in which this kind of instruction is most likely to take place. The literature documents three main relationships: the association between inquiry-oriented teaching and student achievement, the characteristics of students who receive inquiry-oriented instruction, and the characteristics of teachers who use inquiry-oriented instruction. We review each of these in turn.

Inquiry-Oriented Teaching and Student Achievement

Although some research suggests that students exposed to inquiry-oriented instruction obtain a deeper understanding of science (Brown & Campione, 1994; Cognition and Technology Group at Vanderbilt, 1992) and exhibit higher achievement (Glasson, 1989; Mechling & Oliver, 1983; Shymansky et al., 1983; Tal, Krajcik, & Blumenfeld, 2006; Von Secker & Lissitz, 1999), much of the evidence is based on small-scale qualitative or correlational studies. In summarizing the results of these and other studies, Anderson (2002) interpreted the findings as suggesting that inquiry-oriented teaching "can work" (p. 4). For example, analyzing data from the 1990 High School Effectiveness Study, conducted by the National Center for Education Statistics (NCES), Von Secker and Lissitz (1999) estimated relationships between teachers' self-reports of instructional practices recommended by the NSES and student achievement. Students in this nationally representative sample of 10th graders took tests developed to measure their higher order thinking in science as well as their understanding of fundamental concepts and mastery of basic skills in biology, chemistry, earth science, and physics. They found that traditional science teaching methods were negatively associated with students' science achievement. Mean science achievement increased almost 0.4 *SD*s for every 1-*SD* increase in the amount of emphasis placed on lab experiences that actively engaged students' knowledge of scientific inquiry and literacy. For this study, lab experience was a composite variable measuring how often experiments were done, how often reports on experiments were written, time spent conducting lab periods, emphasis on lab techniques, and the use of working in small groups (Von Secker & Lissitz, 1999). Glasson (1989) found similar results in a study that randomly assigned 54 9th graders to physical science classrooms that implemented either a hands-on experiment guided by students or a teacher demonstration laboratory method. Students in the hands-on laboratory classroom scored higher (approximately 0.5 *SD*s) on a test of procedural knowledge compared with students in the teacher demonstration class. Earlier meta-analyses published in the *Journal of Research in Science Teaching* are consistent with these results. For example, Shymansky et al. (1983) found that on the basis of 105 studies, 63% of students who experienced a "new science curriculum"—as defined by an emphasis on the nature and process of science, the integration of laboratory activities in the curriculum, and higher cognitive abilities—outperformed those who experienced a more traditional curriculum that emphasized knowledge of scientific facts, laws, theories, and applications, although in many cases, the effects were modest and there was little correlation between positive results and expert ratings of the degree of inquiry in the materials. Concurrently, Wise and Okey (1983) conducted a meta-analysis of 160 studies and coded various teaching strategies. Inquiry-discovery instruction that included inquiry lessons, guided discoveries, and inductive laboratories constituted 15% of the sample; of these studies, cognitive outcomes increased by 0.87 *SD*s for every 1-*SD* increase in the use of inquiry-discovery teaching. In contrast, Lott (1983) found only small differences between "inductive" (which incorporates aspects of what the NSES considers inquiry) and "deductive" approaches, although the findings were in favor of the inductive approach. Finally, Rosebery, Warren, and Conant (1992) found that inquiry-oriented strategies enhanced scientific ways of thinking, talking, and writing for English-language learners and helped them acquire English and reasoning skills. Although these studies suggest a relationship between inquiry-oriented instruction and improvements in attitudes toward science and in student achievement, the evidence that there is a causal link between these variables is considerably weaker; additionally, the evidence does not point to specific subjects, topics, or grades in which inquiry-oriented methods are most important. Furthermore, other studies suggest that too much inquiry or "discovery learning" (e.g., no teacher intervention beyond the suggestion of a learning objective) can be less effective than direct or teacher-centered instruction (Klahr & Nigam, 2004).

Who Gets Inquiry-Oriented Instruction?

Although students in high-poverty schools appear less likely to encounter inquiry-oriented teaching than students in low-poverty schools, classes with high-achieving students may also be less likely to be exposed to these instructional strategies. Results from a survey of teachers and principals participating in the National Science Foundation's (NSF) Local Systemic Change Through Teacher Enhancement (LSC) initiative (conducted in 24 communities across the country) found that among 300 randomly selected teachers in K–8 schools, those with high proportions of students receiving free or reduced-price lunch were less likely to use inquiry- or reform-based teaching practices or implement investigative culture classrooms than those with more advantaged student bodies (Supovitz & Turner, 2000). In contrast, Smerdon, Burkam, and Lee (1999), analyzing a nationally representative sample of 3,660 tenth-grade students, found that higher achieving students received more didactic or teacher-centered instructional practices than their peers. These findings are contrary to conventional wisdom suggesting that inquiry-based teaching practices are more often used with high-achieving students. Breaking the data down by high school subject can help in the interpretation of these findings. Students in both chemistry and physical science classrooms (the first of which tends to enroll very high achieving and high-socioeconomic-status 10th graders and the second of which tends to enroll lower achieving and low-socioeconomic-status 10th graders) spent more time working on laboratory projects and had greater access to lab equipment. In effect, students who were exposed to more inquiry instruction were more likely to be enrolled in classes with more lab time and less "whole-class instruction." These findings suggest that inquiry-based instructional practices are more likely to be present in science courses with lab components regardless of student ability but that overall, low-achieving students receive more inquiry-based instruction. Von Secker (2002) noted these conflicting findings and concluded that although inquiry-based teaching practices are associated with increases in the achievement of all students, the full implementation of inquiry-oriented standards could affect different students in different ways, potentially exacerbating existing inequities in critical thinking skills while decreasing gaps in other areas, such as laboratory skills.

What Teacher Background Characteristics Are Associated With the Implementation of Inquiry-Oriented Teaching?

Prior research has shown mixed results regarding how teachers' experience level influences their use of inquiry-oriented instruction. Furthermore, there has been relatively little research on how science teachers' content knowledge is related to the content of their instruction, although the relationship is positive in the field of mathematics and case study work in science suggests its importance.

Correlational analyses suggest that the number of years of teaching experience is associated with improved student performance, although the impact of experience appears to level off after about 3 years (Darling-Hammond, 2000; Murnane & Phillips, 1981; Rivkin, Hanushek, & Kain, 1998; Mayer, Mullens, & Moore, 2001). In their study of 10th-grade science instruction, however, Smerdon et al. (1999) found that students taught by teachers with less teaching experience received more inquiry-led instruction than students taught by more experienced teachers. In contrast, in a study of 14 middle and high school science teachers, Luft (2001) found that beginning (less than 2 years of teaching experience) science teachers were less likely to implement extended cycles of inquiry-oriented teaching practices compared with their more experienced colleagues. The relationship between experience and the use of reform-oriented curricula is also mixed in mathematics. For example, in a nationally representative study of 8th-grade mathematics teachers, Smith, Desimone, and Ueno (2005) found beginning teachers less likely to emphasize conceptual learning goals but equally likely to use conceptual teaching strategies in class.

Although there is little prior research on the relationship between teachers' content knowledge, certification areas, or subject areas in which they hold degrees and their levels of implementation of inquiry-based instruction in science classes, prior studies, particularly in mathematics, suggest that teachers require strong content knowledge to successfully implement inquiry-oriented instructional strategies (Loucks-Horsley, Hewson, Love, & Stiles, 1998; Ma, 1999; Remillard, 2000; Rowan, Chiang, & Miller, 1997; Schneider & Krajcik, 2002). For mathematics, Smith et al. (2005) found that preparedness to teach mathematics content—measured by certification in math, holding a degree in mathematics, and teachers' self-reports of their preparedness to teach particular science topics—were associated with reform-oriented teaching, measured as increased emphasis on conceptual learning goals for students and the increased use of reform-oriented teaching strategies. Although it has been argued that inquiry-guided instruction in science also requires teachers to obtain greater knowledge and understanding of science as inquiry and to redefine their views on the purpose of education (Anderson, 2002), the link between teacher content knowledge and use of inquiry in teaching is not well established in the literature.

To What Extent Is Participation in Professional Development Associated With Increased Use of Inquiry Teaching?

Professional development is a practical approach to providing the current science teacher workforce with knowledge pertaining to inquiry-oriented teaching practices (Abd-el-khalick, Bell, & Lederman, 1998; Luft, 2001; Supovitz & Turner, 2000). Many teachers participate in professional development to improve their teaching skills and knowledge. Although it can take many forms, including study groups, curricular development committees, and networking (Loucks-Horsley et al., 1998), the most common forms of professional development are workshops, seminars, and formal college or university courses (Garet, Birman, Porter, Yoon, & Desimone, 2001). Furthermore, although there are many characteristics of professional development that contribute to its effectiveness, such as active learning and coherence with other reforms (e.g., Loucks-Horsley et al., 1998), two of the most important characteristics are a focus on subject-matter content and the amount of time spent on the professional development activities (Cohen & Hill, 2000; Desimone, Porter, Garet, Suk Yoon, & Birman, 2002; Garet et al., 2001; Newmann, Smith, Allensworth, & Bryk, 2001).

Prior research suggests that participation in sustained inquiry-based or problem-solving-focused professional development is associated with an increased use of inquiry in science classes. For example, in a study of seven teachers in Grades 3–6, Luft and Pizzini (1998) found that teachers who completed a yearlong in-service professional development program centered on problem-solving methodology provided their students with more opportunities to design questions and implement strategies to answer scientific questions. In a follow-up study, Luft (2001) showed that teachers who participated in a yearlong Search, Solve, Create and Share problem-solving model in-service program increased their use of inquiry-led classroom discussions. A study by Abd-el-khalick et al. (1998) found that teachers who participated in an inquiry-based service program geared toward improving instruction gained an "adequate" understanding of the role of subjectivity and creativity in science. Finally, Supovitz and Turner (2000) found that teachers in districts participating in the NSF's LSC initiative who had no inquiry-based professional development were less likely (0.4 *SD*s) to create an inquiry-oriented classroom compared with teachers who received inquiry-centered professional development. The extent of teachers' implementation of reform-based practices was measured by questions that asked about, among other indicators, the frequency of inquiry-based teaching practices, such as engaging students in hands-on activities, having students design or implement their own investigations, journaling, and working on "extended science investigations or projects" (Supovitz & Turner, 2000). Compared with the average teacher, teachers with 40–79 hours of inquiry-based professional development continued to use more traditional practices. Only after 80 hours of inquiry-based professional development did teachers report using inquiry-based teaching practices more often (0.2 *SD*s) than the average teacher. More recently, Banilower, Heck, and Weiss (2007), analyzing longitudinal data from 42 projects participating in the NSF's LSC initiative, found that teachers with the mean number of hours of LSC professional development (which was situated in classroom practice, emphasized the content teachers need to implement the designated instructional materials, and was sustained over time) scored 0.43 *SD*s higher on an investigative culture composite and 0.52 *SD*s higher on an investigative practices composite compared with teachers who had not yet participated in LSC professional development. As these authors pointed out, the districts studied were atypical in that they had unusually strong leadership in science education and were systematically engaged in reforming science instruction.

In addition to links between inquiry-oriented professional development and inquiry-oriented teaching, the literature suggests that inquiry-oriented teaching is also fostered by increased subject-matter content knowledge, which is obtained through content-focused professional development. Previous research has shown that procedural, didactic instruction and the use of lower level cognitive demands are more likely when teachers do not have a strong grounding in the subject matter they are teaching (Carlsen, 1990; Cochran & Jones, 1998; Hashweh, 1987; Putnam & Borko, 1997; Tobin & Tippins, 1993). Work over the past decade suggests that professional development focused on subject-matter content may be an especially important element in changing teaching practice toward more higher order, challenging instruction (e.g., Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Corcoran, 1995; Fennema et al., 1996; Hiebert et al., 1996; Simon & Schifter, 1991; Smith et al., 2005). This is based in part on findings that many teachers lack strong content-specific teaching skills (Reynolds, 1995; Rhine, 1998), especially in science and mathematics.

Along these lines, in the subject of mathematics, Desimone et al. (2002) followed teachers over 3 years and found that after participating in professional development in mathematics for a moderate numbers of hours, teachers significantly increased their use of higher order conceptual instruction in the classroom. Cohen and Hill (2001) found similar links between content-focused professional development and higher order teaching practices in mathematics. The existence of analogous relationships between science-focused professional development and inquiry-oriented instruction are plausible but have been less studied, and the evidence that does exist is contradictory. For example, Shymansky et al. (1983), in a meta-analysis of 105 experimental studies, found that students who were exposed to new science curricula (defined as developed after 1955, emphasizing the nature and processes of science, integrating laboratory activities as an integral part of the class routine, and emphasizing higher cognitive abilities and appreciation of science) in instances in which teachers had received in-service training did not perform as well as students whose teachers had received no in-service training. These counterintuitive results are less significant when one considers that only 30% of the studies in the meta-analysis reported any data regarding teachers' in-service backgrounds. Additionally, the characteristics of in-service training were not categorized or qualified in any way.

Our Study

As states and districts refocus their reform efforts to include student achievement in science, it is important to know whether the two policy levers most associated with improving teacher quality—reducing out-of-field teaching and increasing teachers' professional development opportunities—are associated with the amount of inquiry-oriented instruction students are exposed to. Although there have been a few quantitative studies looking at the relationship between teacher characteristics and the use of inquiry-oriented teaching in science, none of the studies reviewed above examined how teachers' content knowledge influenced their use of inquiry in teaching or whether teachers with different levels of content knowledge were differentially influenced by their participation in content-related professional development activities. In our analysis, we examined the relationship between these two policy-malleable variables and the instructional goals and strategies teachers used in eighth-grade science classes. Specifically, our study focused on two main research questions. First, we asked, How are teacher credentials—degree major, certification, and years of experience—related to teachers' use of different instructional strategies? On the basis of research that links credentials in mathematics and student achievement in mathematics, as well as the underlying assumption of NCLB that having a degree in science leads to higher quality instruction in science, we tested the hypothesis that teachers with (a) bachelor's or higher degrees in science, (b) regular teaching credentials and/or a certifications in science, and/or (c) more years of experience would be more likely to emphasize conceptual objectives and use reform-oriented activities in their science classes compared to their newer colleagues without degrees in science and without certification.

Second, we asked, To what extent does participation in content-related professional development mediate the relationships between teacher credentials and teachers' use of different instructional strategies? We expected that teachers' participation in professional development activities focused on science content would increase their content knowledge and thus their likelihood of using inquiry-based instruction in the classroom. This hypothesis is grounded in the studies discussed above that link content-focused professional development with conceptual or higher order instruction. Our hypothesis is consistent with the notion that deepening teachers' content knowledge in the subject they are teaching increases their ability to use multiple teaching strategies and more inventive, challenging instruction (Ball, 1990; Sykes, 1996).

Because professional development participation has become one of the primary ways that veteran teachers can attain "highly qualified" status in many states, it is important to know whether increased participation is likely to lead to increased use of reform-oriented teaching in the classroom. As part of NCLB funding, the U.S. Department of Education (2007a) allocated $2.85 billion in 2002 and $2.93 billion in 2003 through Improving Teacher Quality State Grants to help teachers become highly qualified if they are not already. Similar levels of funds were budgeted for 2004, 2005, and 2006. Accordingly, if professional development participation is a reasonable indicator of efforts to increase "teacher quality," we would expect that teachers' participation in professional development activities focused on science content would predict conceptual objectives and the use of hands-on activities in the science classroom, over and above what teaching experience and holding science-related credentials predict.

A model for how we see these relationships operating is presented in Figure 1. We suggest a sequence of events that fits with previous research, acknowledging that credentials, though easy to examine and monitor, are a weak measure of a teacher's true content knowledge in science.

Method

Data

To examine the relationship between science teachers' credentials, participation in professional development, and instruction, we compiled data from the teacher questionnaire administered as part of the 2000 NAEP Science Assessment. The NAEP has 4th, 8th, and 12th grade national samples; for the current analyses, we used the 8th grade national sample (U.S. Department of Education, 1999, 2000). The national NAEP study is based on a stratified national probability sample of approximately 16,000 eighth graders and their science teachers at 744 schools. NAEP is one of the few nationally representative data sets that surveys teachers about their educational backgrounds, participation in content-related professional development, and use of a wide range of teaching strategies. Recent data collected on nationally representative samples of teachers, such as the NCES Schools and Staffing Survey, do not contain information on instruction. Although the NAEP

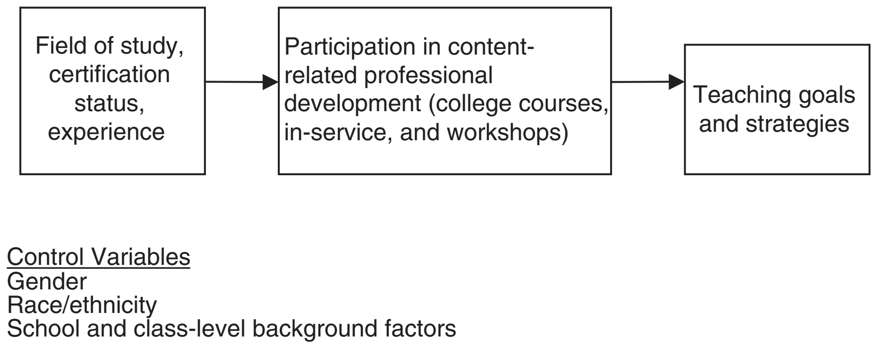

FIGURE 1. *Predicting teachers' use of different teaching strategies on the basis of their credentials, experience, and participation in professional development and controlling for their sex and race and for school factors.*

sample was not specifically designed to estimate the attributes of the teacher population, by using responses provided by teachers on the classes from which students were sampled, we could examine relationships between the teaching strategies that particular 8th grade science teachers reported using and a number of teacher characteristics that are of interest to policy makers.

Sample

The 2000 NAEP national sample was chosen using a complex multistage design. First, schools were sampled from each geographic area across the country, and students were sampled from each school selected (U.S. Department of Education, 2003). Certain schools were oversampled to ensure that the data contained sufficient numbers of these schools, which were in otherwise small categories.[1] Although the teacher data in NAEP were not drawn as a probability sample from all middle school science teachers, they are the only large-scale data currently available that contain information on science teachers' instructional practices, their educational backgrounds, and their levels of participation in content-focused professional development. The limitations of the sample are that teachers from small schools have a lesser chance of being selected (because sampling is done with probability proportional to student enrollment) and that teachers of smaller groups (such as English-language learners or special education students) also have a smaller probability of being selected.

The NAEP data are organized in student-level files, and information on classrooms, teachers, and schools is available only for those teachers who are matched to sampled students who completed the NAEP assessment. We created a teacher-level data file by reaggregating teacher reports that had been disaggregated to the student level in the data file made available by NCES.[2] Using a similar strategy, we created a school-level data file. The final operational sample consisted of 1,073 teachers at 593 schools.

To make our findings more closely representative of the teachers of eighth grade science students sampled in NAEP, we adjusted for oversampling of schools by using school-level sampling weights in the analyses. Because between 20 and 25 students were randomly sampled from all schools, the probability that every eighth grade science teacher taught at least 1 of the eighth grade students sampled was high (i.e., sampled students were likely to be distributed across all of the eighth grade science classes). To simplify the description below, we describe this sample as "eighth grade science teachers," although there were some teachers from this broader population who might have been excluded from our sample (e.g., teachers for whom all of the students whom they taught science were excluded from the NAEP sampling frame, including a small percentage of students whom the school deemed ineligible for the assessment because of very limited English proficiency or disabilities that inhibited testing). Because few eighth grade science teachers were likely to have taught only students who were excluded from the NAEP, we generalized our results to eighth grade science teachers, although given the caveats noted above, we encourage the replication of our results with a nationally representative sample of science teachers.

Measures

Below, we briefly describe how we measured the following variables: (a) instruction, (b) teacher credentials, (c) professional development participation, and (d) control variables. Table 1 provides the exact NAEP questions that correspond to each of the variables, how we coded the variables, and their means and standard deviations as applicable.

Instruction

This study relied on teacher self-reports of instructional activities. Research has shown that survey measures of teaching, especially composite measures such as those we used in this study, can be effective in describing and distinguishing among different types of teaching practices (Mayer, 1999). Self-report surveys, however, tend not to be good instruments for measuring affective aspects of instruction, such as warmth and student-teacher interaction and engagement (Burstein et al., 1995).

Several studies have shown that teachers' self-reports of their own teaching on anonymous sample surveys are moderately to highly correlated with classroom observations and teacher logs (correlations average between .7 and .8) and that one-time surveys asking teachers questions about the content and strategies they emphasize can be valid and reliable in measuring their instruction (Mullens, 1995; Mullens & Gayler, 1999; Mullens & Kasprzyk, 1996, 1999; Schmidt, McKnight, & Raizen, 1997; Shavelson, Webb, & Burstein, 1986; Smithson & Porter, 1994). The risk for social desirability bias is also always a potential problem in survey research, although this risk tends to be highest when the survey is not confidential and/or is used for evaluating teachers (Mayer, 1999). In confidential, nonevaluative circumstances, research has shown that data collected in written, self-administered modes can be as accurate as interviews (Aquilino, 1994, 1998; Aquilino & LoSciuto, 1990; Dillman & Tarnai, 1991; Fowler, Roman, & Di, 1998; Hochstim, 1967). Given that NAEP is anonymous, confidential, and never linked back to teachers and focuses on behavioral rather than relational or affective instructional practices, we have a reasonable level of confidence in the self-report items as indicators of teachers' instruction on the limited dimensions in our study. Still, cross-sectional survey data are not sufficient evidence for establishing causal relationships, although they can be suggestive and helpful in establishing the range of effect sizes that might be detectible in a randomized field trial (e.g., Rowan, Correnti, & Miller, 2002).

We used the reform literature on science and the NSES to guide our choice of measures of instruction from the NAEP teacher questionnaire. NAEP asks how often teachers use different instructional strategies or emphasize different conceptual objectives. The items fall into four broad categories. Three of these categories are related to the frequency of the use of procedural activities (use of multiple-choice tests, assignments to read a science textbook, science tests or quizzes, and emphasis on science facts and terminology), reporting and writing activities (individual or group projects that take a week or more, short and long written responses, using lab notebooks or journals, oral science reports, and written science reports), and hands-on activities in class (evaluate students on the basis of hands-on activities, have students work together on activities, do hands-on activities or investigations in science, talk about measurements and the results of students' hands-on activities, and emphasize developing lab skills). The fourth category is related to the amount of emphasis that teachers give to conceptual objectives (developing science problem-solving skills, learning about the relevance of science to society and technology, knowing how to communicate ideas in science effectively, developing students' interest in science, and developing data analysis skills). Previous research on "higher order" or conceptual teaching has called for the separation of conceptual goals and strategies, because they are likely to reflect different dimensions of instruction (Raudenbush, Rowan, & Cheong, 1993). The NAEP data allowed such a distinction, so our initial plan was to construct separate measures of activities and objectives.

Using factor analysis (principal components with varimax rotation), we assessed the degree to which 19 teacher-level questions asked in the NAEP teacher survey (detailed in Table 1 and described below) tap these four different dimensions of instruction (frequency of the use of procedural activities, reporting and writing activities, and hands-on activities in class and the amount of emphasis teachers give to conceptual objectives).[3] Our initial plan was to combine dimensions that were not empirically distinguishable. The 19 items loaded on four factors with eigenvalues greater than 1, cumulatively explaining 53% of the variance across the items. These four factors mapped well to the dimensions of frequency of the use of procedural activities, reporting and writing activities, hands-on activities in class and the amount of emphasis teachers give to conceptual objectives, with 4 items loading on procedural teaching activities (loadings between .57 and .72), 5 items loading on reporting and writing activities (loadings between .32 and .73), 5 items loading on hands-on activities (loadings of .69 and .87), and 5 items loading on emphasis on conceptual goals (factor loadings between .53 and .69). We then developed composites and indexes on the basis of the sum of these groups of items, reported below. Following the practices of Raudenbush et al. (1993), each of the composites was standardized to have a mean of 50 and a standard deviation of 10 to aid interpretation of the analyses (standardized and unstandardized values are included in Table 1).

Procedural activities was an index composed of the average of four items measuring how often teachers reported using the following to teach and assess students in science (on a 4-point scale ranging from *never or hardly ever* to *almost every day*): (a) multiple-choice tests, (b) assignments to read a science textbook, (c) science tests or quizzes, and (d) an emphasis on science facts and terminology ($\alpha = .63$).

Reporting and writing activities was an index composed of the average of five items measuring how often teachers reported using the following to teach and assess students in science (on a 4-point scale ranging from *never or hardly ever* to *almost every day*): (a) individual or group projects that take a week or more, (b) short and long written responses, (c) lab notebooks or journals, (d) oral science reports, and (e) written science reports ($\alpha = .59$).

Hands-on activities was an index composed of the average of five items measuring how often teachers reported doing each of the following (on a 4-point scale ranging from *never or hardly ever* to *almost every day*): (a) evaluating students on the basis of hands-on activities, (b) having students work together on activities, (c) doing hands-on activities or investigations in science, (d) talking about measurements and the results of students' hands-on activities, and (e) emphasizing developing lab skills ($\alpha = .84$).

Conceptual emphasis was an index composed of the average of five items measuring how much emphasis teachers reported giving each of the following objective for their students (on a 3-point scale ranging from *little or no emphasis* to *heavy emphasis*): (a) developing science problem-solving skills, (b) learning about the relevance of science to society and technology, (c) knowing how to communicate ideas in science effectively, (d) developing students' interest in science, and (e) developing data analysis skills ($\alpha = .71$).

The correlation between these indexes was modest, ranging from –.20 between the use of procedural activities and the use of hands-on activities to .48 between the use of reporting and writing activities and the use of hands-on activities. In general, there was a low or negative relationship between the use of procedural activities and reform activities and goals. The correlation among reform activities and goals ranged from .43 to .48 (see Table 2).

Although our measures do not encompass all of the dimensions of the NSES in their depth and complexity, they are consistent and aligned with key ideas. We expected teachers who emphasized conceptual objectives to be more likely to treat learning as an active process and to organize lessons that involved "making observations; posing questions; . . . planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and

TABLE 1
Measures of Key Variables and Descriptive Statistics

Variable	Unweighted *M*	Unweighted *SD*	Weighted *M*	Weighted *SE*
Dependent variables (teaching quality)				
Procedural Activities (sum of 4 items;	49.97	9.91	50.51	0.41
α = .63, standardized)[a]	*2.75*	*0.467*	*2.780*	*0.019*
Reporting and writing activities	49.86	9.93	49.29	0.47
(sum of 5 items; α = .59, standardized)[b]	*1.98*	*0.440*	*1.96*	*0.020*
Hands-on activities (sum of 5 items;	49.89	10.01	49.26	0.44
α = .84, standardized)[c]	*2.52*	*0.590*	*2.48*	*0.026*
Conceptual emphasis (sum of 5 items;	50.00	10.02	49.41	0.43
α = .71, standardized)[d]	*2.41*	*0.402*	*2.39*	*0.017*
Teacher-level independent variables				
Sex[e]	0.58		0.57	0.02
Race and ethnicity[f] (recoded)				
White	0.84		0.87	0.01
Black	0.10		0.08	0.01
Hispanic	0.04		0.02	0.01
Asian	0.02		0.01	0.00
Other	0.01		0.01	0.00
Years of science teaching experience[g] (recoded)				
2 years or less	0.20		0.20	0.02
3–5 years	0.21		0.20	0.02
6–10 years	0.20		0.22	0.02
11 years or more	0.39		0.38	0.02
Certificate[h] (recoded)				
Full certification (including advanced professional, regular/standard, probationary)	0.82		0.86	0.01
Partial certification (including temporary, provisional, or emergency state certificate)	0.08		0.07	0.01
No state certification (including certificate not from the state and no certificate)	0.10		0.08	0.01
Certificate in science[i] (recoded)				
0 = no teaching certification or certification not in science (or not known)	0.40		0.39	0.02
1 = science certification	0.60		0.61	0.02
1 = certification unknown; 0 = certification known	0.12		0.11	0.02
Degree-major[j] (recoded)				
Science graduate major	0.12		0.10	0.01
Science undergraduate major	0.34		0.33	0.02
Science minor	0.08		0.08	0.01
Science education major or minor	0.08		0.08	0.01
Other science (degree in science but misspecification or unknown degree type	0.01		0.01	0.00
No science or science education major/minor	0.38		0.40	0.02
Any science degree	0.62		0.60	0.02
Any education major/minor	0.87		0.89	0.01
Professional development workshop participation[k] (recoded)				
0 hours of workshop/seminars	0.12		0.13	0.02
Less than 6 hours of workshops/seminars	0.21		0.20	0.02

(continued)

TABLE 1 (continued)

Variable	Unweighted *M*	Unweighted *SD*	Weighted *M*	Weighted *SE*
Between 6 and 15 hours of workshops/seminars	0.29		0.27	0.02
Between 16 and 35 hours of workshops/seminars	0.19		0.22	0.02
More than 35 hours of workshops/seminars	0.20		0.18	0.01
As a continuous variable, using the median values the categories of response	16.50		16.17	0.57
Professional development class participation[l]	0.81		0.74	0.05
School-level independent variables				
School type[m] (recoded)				
Regular school	0.67		0.76	0.02
Magnet school (or regular school with a magnet program)	0.06		0.05	0.01
Special education or alternative (nontraditional) school	0.01		0.00	0.00
Private (religious or independent)	0.26		0.18	0.02
School socioeconomic level (percentage of students eligible for free lunch)[n]				
(Recoded as a continuous variable, using the median values of the categories of response: 0%, 1%–5% (coded 3), 6%–10% (coded 8), 11%–25% (coded 18), 26%–50% (coded 38), 51%–75% (coded 63), 76%–99% (coded 87.5), and 100%)	34.11	30.23	35.89	1.82

Note. Sample size = 1,073 teachers. Standard deviations omitted for dichotomous variables. Unstandardized values for dependent variables (teaching quality) are in italics.

a. "How often do you use each of the following to teach/assess students in science? (a) multiple-choice tests; (b) assignments to read a science textbook; (c) science test/quiz; and (d) emphasize science facts/terminology" (1 = *never or hardly ever*; 2 = *once or twice a month*, 3 = *once or twice a week*, 4 = *almost every day*).

b. "About how often do you do each of the following? (a) individual/group projects that take a week or more; (b) short/long written responses; (c) using lab notebook/journal; (d) oral science reports; and (e) written science reports" (1 = *never or hardly ever*; 2 = *once or twice a month*, 3 = *once or twice a week*, 4 = *almost every day*).

c. "About how often do you do each of the following? (a) evaluate students based on hands-on activities; (b) have students work together on activities; (c) do hands-on activities or investigations in science; (d) talk about measurements and results of students' hands-on activities; and (e) emphasize developing lab skills" (1 = *never or hardly ever*; 2 = *once or twice a month*, 3 = *once or twice a week*, 4 = *almost every day*).

d. "About how much emphasis will you give to each of the following objectives for your students? (a) developing science problem-solving skills; (b) learning about the relevance of science to society and technology; (c) knowing how to communicate ideas in science effectively; (d) developing students' interest in science; and (e) developing data analysis skills" (1 = *little or no emphasis*; 2 = *moderate emphasis*; 3 = *heavy emphasis*).

e. "What is your gender?" (0 = *male*; 1 = *female*).

f. "Which best describes you?"

g. "Counting this year, how many years in total have you taught science?"

h. "What type of teaching certificate do you have in this state in your main assignment field?"

i. "Do you have teaching certification in any of the following areas that is recognized by the state in which you teach? Elementary Science, Middle/Junior High School, or Secondary Science"

j. "What is the highest academic degree you hold?" "What were your undergraduate major fields of study?" "What were your graduate major fields of study?" and "What were your undergraduate and graduate minor fields of study?"

k. "During the last year, how much time in total have you spent in professional development workshops or seminars in science or science education?"

l. "During the last 2 years, how many college or university courses have you taken in science or science education?"

m. "What type of school is this?"

n. "Does your school participate in the National School Lunch Program?" and "During this school year, about what percentage of students in your school was eligible to receive a free or reduced-price lunch through the National School Lunch Program?"

TABLE 2
Correlation of Instructional Variables

	Procedural activities	Reporting and writing activities	Hands-on activities	Conceptual emphasis
Procedural activities	1.00			
Reporting and writing activities	.05	1.00		
Hands-on activities	–.20	.48	1.00	
Conceptual emphasis	1.00	.43	.46	1.00

communicating the results" (National Research Council, 1996, p. 23). We also expected the index of hands-on activities to be greater in classrooms that emphasized inquiry more, although some of the activities may be common in classrooms in which science kits or experiments were used for demonstration rather than exploration. Reporting and writing activities could also be prevalent in a traditional classroom, although these types of activities, as suggested in the NSES, are likely to make an inquiry-focused classroom more effective. Although only the indexes reflecting teachers' use of hands-on activities and their emphasis on conceptual objectives obtained conventional standards of reliability ($\alpha > .70$), we chose to examine all four measures as dependent variables to assess patterns across the three constructs most closely aligned with the NSES and to have the use of the procedural activities index as a contrast.

Teacher Certification

Although the criteria for becoming certified to teach vary across states, certification, either through traditional or alternative programs, is a requirement for teachers to achieve highly qualified status under NCLB. Thus, although the amount of information about science content knowledge or pedagogical preparation that certification might signal differs across states, we included certification in our analysis to assess its association, on average, with science teachers' instruction.

In the NAEP, teachers answered questions regarding their overall certification status and whether their certification was in (a) elementary or middle or junior high school general education, (b) elementary science, or (c) middle or junior high school or secondary science. Using the first question, we created a series of dichotomous variables: (a) full certification, which included teachers who reported holding advanced professional certificates, regular or standard state certificates, or probationary certificates in their main assignment field;[4] (b) temporary certification, which included temporary, provisional, or emergency state certification in teachers' main assignment fields (asked all in the same question in NAEP); and (b) no state certification, which included teachers who responded that they did not have certification in their main assignment fields or that their certification was from accrediting bodies other than states. Using the second question, we created an additional dichotomous variable indicating whether a teacher's certification was in elementary, middle or junior high, or secondary science (science certification). From the response patterns in the NAEP questionnaire, it is not always possible to tell whether teachers had certification in science (e.g., they answered neither "yes" nor "no" to the questions). We coded these respondents as having no science certification but included a dummy variable to indicate that data on this item were missing (science certification unknown).

Major Fields of Study

Teachers were also asked to report the fields of their undergraduate and graduate majors and minors. Because teachers could report multiple degrees, we created a set of six mutually exclusive categories to hierarchically classify teachers with varying levels of formal education in science. We considered that teachers who majored in science in their graduate programs represented teachers with the most advanced education in the field (science graduate major;

12% of the sample). Teachers with the second strongest science education were those who majored in science in their undergraduate programs, but not in their graduate programs, because these are mutually exclusive categories (science undergraduate major; 34% of the sample). We defined teachers with the third strongest science education as those who had minors in science (science minor; 8%) or science education majors or minors (science education major or minor; 8%). The fifth category included teachers who indicated that they had degrees in science but for whom the degree levels were unknown (other science; 1%). Finally, the sixth category included teachers with no major or minor in either science or science education at either the undergraduate or graduate levels (no science or science education major or minor; 38%). We also created an additional dichotomous variable indicating whether a teacher had a major or minor in any education field at any level (any education degree). Thus, if a teacher had both a bachelor's degree in physics and a bachelor's degree in secondary education, for example, he or she would be coded as 1 on science undergraduate major, 1 on any education degree, and 0 on the other variables. If a teacher had a bachelor's degree in science education and no higher degree in science, he or she would be coded as 1 on science education major or minor, 1 on any education major, and 0 on the other variables.

Professional Development Participation

The NAEP questionnaire asked teachers separately about the number of college or university courses they had taken in the past 2 years in science or science education and the number of hours they spent in professional development workshops or seminars in science or science education during the prior year. We coded both of these on a 5-point scale, with the first (college courses) coded as an integer ranging from 0–4 and the second (professional development hours) coded as the midpoint of the five categories included in the questionnaire item: (a) none (coded 0), (b) less than 6 hours (coded 3), (c) 6–15 hours (coded 10.5), (d) 16–35 hours (coded 25.5), and (e) more than 35 hours (coded 40). These measures included only content focus (science or science education) and duration as dimensions of participation. Having a measure of the span of time over which the activity was spread (not available in the NAEP) would have given us a more complete measure of duration, but a high number of contact hours has been shown to be associated with effective professional development (Garet et al., 2001). In addition, the NAEP teacher questionnaire does not include other dimensions of professional development related to effectiveness (see Desimone et al., 2002; Garet et al., 2001), including (a) the form of the activity (i.e., whether it is a reform activity, such as a study group or network, in contrast to a traditional workshop or conference); (b) the degree to which the activity emphasizes the collective participation of groups of teachers from the same school, department, or grade level, as opposed to the participation of individual teachers from many schools; (c) the extent to which the activity offers opportunities for teachers to participate in active learning, by reviewing student work, for example, or through obtaining feedback on their teaching; or (d) the degree to which the activity promotes coherence in teachers' professional development by incorporating experiences that are consistent with teachers' goals, aligned with state standards and assessments, and encourage continuing professional communication among teachers. Thus, we expected effect sizes associated with this variable to be lower than those reported in studies of professional development that more closely align to these additional dimensions of quality.

Control Variables

Additional characteristics of teachers and schools may have been associated with our key independent variables and may also have predicted the teachers' use of different instructional practices. Thus, we statistically controlled for these factors to reduce the potential bias of estimates of the relationships between our key independent variables and teaching quality.

Teacher-level control variables included sex (*female* = 1, *male* = 0), race and ethnicity (White non-Hispanic as the reference category, Black non-Hispanic, Asian, and other), and years of teaching science (2 years or less, 3–5 years, 6–10 years as the reference category, and 11 and more years).[5]

For school characteristics, we controlled for school type and socioeconomic level. School type included four categories: regular as the reference category, magnet, special education or alternative, and private. Because the socioeconomic level of the student body was not directly measured in NAEP, we used as a proxy measure the percentage of students who were eligible for free or reduced-price lunch through the National School Lunch Program. The variable included eight categories, which we recoded using the midpoint of each category: 0%, 1%–5% (coded 3), 6%–10% (coded 8), 11%–25% (coded 18), 26%–50% (coded as 38), 51%–75% (coded 63), 76%–99% (coded 87.5), and 100%.

Analysis

The presentation of our analysis is organized around our two research questions. To answer our first question, which examined the relationship between credentials and teaching, we used a two-level hierarchical linear model to predict each of four types of instruction (frequency of the use of procedural activities, reporting and writing activities, and hands-on activities in class and the amount of emphasis that teachers gave to conceptual objectives) as a function of experience teaching, certification status, and major field of study (Level 1 variables), while holding constant teacher background characteristics (Level 1) and school characteristics (Level 2). We then added participation in content-related professional development to these models to examine the degree to which the relationships between credentials and experience on teachers' emphasis on different instructional activities and objectives were mediated through professional development participation. Finally, we interacted our best measure of content knowledge, whether teachers had degrees in science, with their levels of participation in professional development to assess whether professional development participation had a stronger influence on increasing the use of reform-oriented science activities on teachers who were likely to have weaker content knowledge. Correlations of teachers' experience teaching, certification status, major fields of study, and participation in content-related professional development are listed in Table 3. Among variables that were not just dummy variables derived from a single categorical variable, the highest correlation was .505 (certification in science and having a degree in science), indicating that multicollinearity was not a serious cause for concern when interpreting our key independent variables.

Tables 4 through 7 show the results of four separate two-level hierarchical linear modeling analyses predicting teachers' emphasis on the use of procedural activities (Table 4), use of reporting and writing activities (Table 5), use of hands-on activities (Table 6), and emphasis on conceptual objectives (Table 7), first as a function of teachers' background characteristics, years of experience teaching, certification status, and levels of formal education in science, controlling for teacher background and school-level characteristics (Model 1 in each table). The additive explanatory power of participation in content-focused professional development was then examined in Model 2; interactions between level of content knowledge, proxied here by having a degree in science, and participation in professional development were added in Model 3.

How Are Teachers' Credentials—Degree Major, Certification, and Years of Experience—Related to Their Use of Different Instructional Strategies?

Neither type of certification nor having a major in science or science education was associated with teachers' self-reported use of procedural teaching activities in eighth grade science classes. Inexperienced teachers were more likely to use procedural activities (although this was only significant at the $p < .10$ level), as were teachers with 11 years or more of experience ($\beta = 1.98$, $p = .039$; Table 4, Model 1). Participation in professional development, either measured as hours spent in workshops or number of college or university courses taken, had neither a strong nor a statistically significant relationship with teachers' use of procedural activities (Table 4, Model 2).

(text continues on p. 188)

TABLE 3

Correlations of Key Teacher Variables

	Full certification	Temporary certification	Science certification	Science certification unknown	Science graduate major	Science undergraduate major	Science minor (graduate or undergraduate)	Science education degree	Degree in science at unindicated level	Any science degree	Education degree	Professional development: Workshops (hours)	Professional development: Number of college classes
Full certification	1.000												
Partial certification	–.644	1.000											
	.000***												
Science certification	.365	–.078	1.000										
	.000***	.010**											
Science certification unknown	.056	.052	–.451	1.000									
	.066†	.088†	.000***										
Science graduate major	.034	–.051	.128	–.057	1.000								
	.268	.096†	.000***	.060***									
Science undergraduate major	.006	.023	.291	–.143	–.264	1.000							
	.851	.445†	.000***	.000***	.000***								
Science minor (graduate or undergraduate)	.026	–.011	.112	–.073	–.106	–.204	1.000						
	.404	.719	.000***	.017***	.001***	.000***							
Science education degree	.080	–.049	.141	–.062	–.106	–.204	–.082	1.000					
	.008**	.109	.000***	.042***	.001***	.000***	.007**						
Degree in science at unindicated level	.018	.018	–.009	.002	–.045	–.088	–.035	–.035	1.000				
	.548	.561	.769	.953	.137	.004	.250	.250					

(continued)

TABLE 3 (continued)

	Full certification	Temporary certification	Science certification	Science certification unknown	Science graduate major	Science undergraduate major	Science minor (graduate or undergraduate)	Science education degree	Degree in science at unindicated level	Any science degree	Education degree	Professional development: Workshops (hours)	Professional development: Number of college classes
Science degree	.091	–.040	.505	–.251	.287	.554	.222	.222	.096	1.000			
	.003**	.195	.000***	.000***	.000***	.000***	.000***	.000***	.002**				
Education degree	.387	–.198	.126	.067	–.043	–.049	.007	.111	.025	–.006	1.000		
	.000***	.000	.000***	.027***	.161	.107***	.826	.000***	.412	.839			
Professional development: workshops	.098	.002	.225	–.099	.077	.141	.036	–.016	.072	.219	.003	1.000	
(hours)	.001***	.946	.000***	.001***	.011*	.000***	.239	.599	.018*	.000***	.930		
Professional development:	–.083	.124	.083	–.068	.044	.123	.018	–.020	.042	.159	–.093	.308	1.000
number of college classes	.007**	.000***	.006**	.027**	.153	.000**	.556	.522	.174	.000***	.002**	.000***	

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 4

Hierarchical Linear Models Predicting Teachers' Use of Procedural Activities

Variable	Model 1	Model 2	Model 3
Level 1 (teacher)			
Intercept	52.09 (1.88)***	51.84 (1.89)***	51.98 (1.94)***
Sex (female = 1)	–1.10 (0.69)	–1.08 (0.70)	–1.07 (0.70)
Race (reference = White)			
Black	2.65 (1.54)†	2.74 (1.56)†	2.60 (1.52)†
Hispanic	2.00 (1.81)	2.10 (1.75)	1.96 (1.75)
Asian	–0.03 (2.08)	0.27 (2.08)	0.29 (2.04)
Other	–0.86 (2.65)	–0.75 (2.63)	–0.81 (2.42)
Years of experience (reference = 6–10 years)			
2 years or less	2.38 (1.24)†	2.39 (1.26)†	2.15 (1.24)†
3–5 years	1.73 (1.12)	1.71 (1.12)	1.57 (1.12)
11 years or more	1.98 (0.96)*	1.99 (0.96)*	1.81 (0.98)†
Teaching (reference = no certificate state certification)			
Full certification	–2.09 (1.62)	–1.99 (1.63)	–1.82 (1.68)
Temporary certification	–3.02 (1.87)	–2.92 (1.92)	–2.64 (1.93)
Science certification (reference = no certificate or not in science)			
Science certification	0.81 (1.11)	0.91 (1.11)	0.68 (1.09)
Science certification unknown	–0.58 (1.37)	–0.64 (1.37)	–0.70 (1.36)
Science-related major (reference = no science or science education)			
Science major graduate	–2.18 (1.51)	–1.98 (1.49)	
Science major undergraduate	–1.58 (1.07)	–1.34 (1.07)	
Science minor	–1.53 (1.48)	–1.54 (1.48)	
Science education (major or minor)	–1.70 (1.66)	–1.71 (1.67)	
Other science (science degree but level unclear)	1.97 (2.77)	2.31 (2.73)	
Any science degree			–1.44 (0.95)
Any education degree	–0.22 (1.28)	–0.23 (1.28)	–0.17 (1.29)
Professional development workshop		–0.04 (0.03)	0.01 (0.04)
Interaction with "any science degree"			–0.07 (0.05)
Number of college or university courses		–0.16 (0.29)	–0.14 (0.29)
Level 2 (school)			
School type (reference = public)			
Magnet	–6.70 (2.03)***	–6.64 (1.98)***	–6.61 (1.95)***
Special education	–3.38 (3.76)	–3.54 (3.70)	–3.71 (3.67)
Private	1.67 (1.12)	1.47 (1.12)	1.59 (1.12)
Percentage free and reduced-price lunch	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
Variance component			
Level 1 variance	57.6	58.1	57.6
Level 2 variance	23.7***	22.6***	23.3***
df	588	588	588
χ^2	1137.0	1113.3	1125.6
Deviance	7,681.6	7,677.7	7,677.3
Number of estimated parameters	25	27	24

Note. Sample size = 1,073 teachers. Unstandardized coefficients are shown with robust standard errors in parentheses. Percentage free and reduced-price lunch, professional development workshop, and number of college or university courses are grand mean centered.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

In contrast, teachers who majored in science and participated in more content-focused professional development tended to be more likely than their counterparts to use reporting and writing and hands-on activities and were more likely to emphasize conceptual objectives in their science classes. For example, having a graduate degree in a science field ($\beta = 3.69$, $p = .008$) was associated with a 37% of a standard deviation increase in the use of reporting and writing activities in eighth grade science classes compared with teachers who had no major in science or science education (Table 5, Model 1). Having a science major at the graduate level was associated with a 54.5% of a standard deviation increase in the use of hands-on activities ($\beta = 5.45$, $p = .000$), while having an undergraduate major in a science field ($\beta = 3.81$, $p = .000$) or a major or minor in science education ($\beta = 2.76$, $p = .054$) was associated with between a 28% and 38% of a standard deviation increase in the use of hands-on activities (Table 6, Model 1). Increased emphasis on conceptual objectives was similarly associated with having a degree in science, although the coefficient for having an undergraduate major in science education was not significant, and having either an undergraduate major or a minor in science was significant only at the $p < .10$ level (Table 7, Model 1).

The relationship between certification and reform-oriented instruction in science differed across measures. Although being certified to teach science was associated with increased use of reporting and writing activities ($\beta = 3.22$, $p = .015$), the relationship was weaker and not significant for use of hands-on activities ($\beta = .52$, $p = .591$) or for emphasis on conceptual objectives ($\beta = 1.60$, $p = .218$). Having full state-level certification (as opposed to no certification) was associated with increased use of hands-on activities ($\beta = 4.77$, $p = .003$), as was holding temporary certification ($\beta = 3.08$, $p = .062$), but there was no relationship between certification in general and either the use of reporting and writing activities or emphasis on conceptual objectives.

Model 2 added hours of participation in professional development workshops or seminars in science or science education and the number of college or university courses taken in science or science education to the variables included in Model 1. A 10-hour increase in participation in professional development workshops or seminars in science or science education was associated with an 18% of a standard deviation increase in the use of reporting and writing activities ($\beta = .18$, $p = .000$), a 16% of a standard deviation increase in the use of hands-on activities ($\beta = .16$, $p = .000$), and an 11% of a standard deviation increase in emphasis on conceptual objectives ($\beta = .11$, $p = .000$) (Tables 5, 6, and 7, respectively, Model 2). Furthermore, although taking college or university courses in science or science education was not associated with a strong or statistically significant increase in the use of reporting and writing activities or hands-on activities in science class, each additional course taken was associated with a 7.5% of a standard deviation increase in emphasis on conceptual objectives ($\beta = .75$, $p = .005$). At least some of the relationship between having a degree in science or science education and the use of reporting and writing activities, the use of hands-on activities, and emphasis on conceptual objectives can be explained by teachers' with science degrees taking more content-oriented professional development. This can be seen by the 17%–49% reduction in the size of the statistically significant coefficients on the science major variables between Models 1 and 2, when participation in professional development variables are added (Tables 5–7).

To What Extent Does Participation in Content-Related Professional Development Mediate the Relationships Between Teacher Credentials and Teachers' Use of Different Instructional Strategies?

Model 3 in Tables 4–7 adds an interaction between teachers' having a major or minor in a science field or science education and hours of participation in professional development workshops or seminars in science or science education. A positive interaction would imply that participation in professional development activities is likely to have a stronger impact on increasing teachers' use of inquiry-oriented instruction if they already have strong content knowledge (a catalytic effect); a negative coefficient would tell us that professional development has a stronger impact on teachers with

(text continues on p. 192)

TABLE 5

Hierarchical Linear Models Predicting Teachers' Use of Reporting and Writing Activities

Variable	Model 1	Model 2	Model 3
Level 1 (teacher)			
Intercept	42.03 (2.26)***	42.84 (2.14)***	43.03 (2.13)***
Sex (female = 1)	1.97 (0.91)*	2.00 (0.84)*	1.99 (0.85)*
Race (reference = White)			
Black	5.17 (1.78)**	5.27 (1.72)**	5.40 (1.70)**
Hispanic	3.38 (1.66)*	2.82 (1.72)	2.66 (1.69)
Asian	4.39 (2.53)†	2.91 (2.57)	3.18 (2.62)
Other	–0.50 (2.20)	–1.16 (1.92)	–0.49 (1.73)
Years of experience (reference = 6–10 years)			
2 years or less	–1.22 (1.64)	–0.94 (1.50)	–1.31 (1.54)
3–5 years	–1.03 (1.27)	–1.07 (1.21)	–1.36 (1.21)
11 years or more	–1.48 (1.18)	–2.01 (1.06)	–2.03 (1.05)†
Teaching certificate (reference = no state certification)			
Full certification	2.26 (1.92)	1.88 (1.93)	2.02 (1.88)
Temporary certification	1.09 (2.17)	0.99 (2.16)	1.24 (2.12)
Science certification (reference = no certificate or not in science)			
Science certification	3.22 (1.32)*	3.01 (1.32)*	2.68 (1.34)*
Science certification unknown	2.32 (1.29)†	2.53 (1.28)*	2.41 (1.28)†
Science-related major (reference = no science or science education)			
Science major graduate	3.69 (1.38)**	3.10 (1.35)*	
Science major undergraduate	1.07 (1.12)	0.40 (1.05)	
Science minor	0.59 (1.78)	0.80 (1.80)	
Science education (major or minor)	–0.01 (1.43)	0.20 (1.34)	
Other science (science degree but level unclear)	4.31 (1.64)**	3.46 (1.64)	
Any science degree			0.87 (0.90)
Any education degree	1.28 (1.44)	1.19 (1.33)	1.32 (1.32)
Professional development workshop		0.18 (0.03)***	0.24 (0.05)***
Interaction with "any science degree"			–0.10 (0.06)
Number of college or university courses		–0.28 (0.41)	–0.29 (0.42)
Level 2 (school)			
School type (reference = public)			
Magnet	3.08 (2.44)	2.99 (2.23)	3.12 (2.25)
Special education	–1.87 (3.40)	–1.43 (3.10)	–1.34 (3.03)
Private	1.95 (1.11)†	2.88 (1.09)**	3.10 (1.07)**
Percentage free and reduced-price lunch	–0.02 (0.02)	–0.02 (0.02)	–0.02 (0.02)
Variance component			
Level 1 variance	69.4	63.9	64.1
Level 2 variance	18.1***	19.3***	19.3***
df	588	588	588
χ^2	869.0	923.3	911.3
Deviance	7,793.4	7,730.7	7,733.4
Number of estimated parameters	25	27	24

Note. Sample size = 1,073 teachers. Unstandardized coefficients are shown with robust standard errors in parentheses. Percentage free and reduced-price lunch, professional development workshop, and number of college or university courses are grand mean centered.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 6

Hierarchical Linear Models Predicting Teachers' Use of Hands-On Activities

Variable	Model 1	Model 2	Model 3
Level 1 (teacher)			
Intercept	40.11 (1.80)***	41.05 (1.77)***	41.19 (1.73)***
Sex (female = 1)	3.12 (0.77)***	3.05 (0.71)***	2.93 (0.68)***
Race (reference = White)			
Black	2.06 (1.42)	2.02 (1.30)	2.04 (1.25)
Hispanic	3.39 (1.74)†	2.80 (1.65)†	2.73 (1.58)†
Asian	–0.83 (2.70)	–2.13 (2.77)	–1.72 (2.97)
Other	2.57 (2.56)	1.93 (2.55)	2.44 (2.46)
Years of experience (reference = 6–10 years)			
2 years or less	–1.08 (1.15)	–0.96 (1.05)	–1.08 (1.01)
3–5 years	–0.42 (1.11)	–0.36 (1.03)	–0.39 (1.00)
11 years or more	0.63 (1.07)	0.42 (0.91)	0.47 (0.87)
Teaching certificate (reference = no state certification)			
Full certification	4.77 (1.56)**	4.38 (1.59)**	4.52 (1.52)**
Temporary certification	3.08 (1.65)†	2.86 (1.67)†	3.02 (1.65)†
Science certification (reference = no certificate or not in science)			
Science certification	0.52 (0.97)	0.23 (0.89)	0.12 (0.88)
Science certification unknown	0.65 (1.18)	0.88 (1.13)	0.90 (1.13)
Science-related major (reference = no science or science education)			
Science major graduate	5.45 (1.03)***	4.67 (1.07)***	
Science major undergraduate	3.81 (0.84)***	2.88 (0.81)***	
Science minor	1.37 (1.37)	1.48 (1.30)	
Science education (major or minor)	2.76 (1.43)†	2.83 (1.29)*	
Other science (science degree but level unclear)	1.30 (2.39)	0.08 (2.50)	
Any science degree			2.75 (0.74)***
Any education degree	1.47 (1.04)	1.45 (0.97)	1.51 (0.96)
Professional development workshop		0.16 (0.03)***	0.23 (0.05)***
Interaction with "any science degree"			–0.12 (0.05)*
Number of college or university courses		0.32 (0.25)	0.32 (0.25)
Level 2 (school)			
School type (reference = public)			
Magnet	0.78 (2.54)	0.65 (2.21)	0.87 (2.16)
Special education	–2.68 (3.07)	–2.06 (2.90)	–2.17 (2.90)
Private	–2.89 (1.04)**	–2.03 (1.00)*	–1.79 (0.98)†
Percentage free and reduced-price lunch	–0.06 (0.02)**	–0.05 (0.02)***	–0.05 (0.02)***
Variance component			
Level 1 variance	48.5	44.9	44.7
Level 2 variance	26.3***	24.9***	25.1***
df	588	588	588
χ^2	1316.4	1360.5	1366.2
Deviance	7,562.2	7,485.5	7,484.3
Number of estimated parameters	25	27	24

Note. Sample size = 1,073 teachers. Unstandardized coefficients are shown with robust standard errors in parentheses. Percentage free and reduced-price lunch, professional development workshop, and number of college or university courses are grand mean centered.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 7

Hierarchical Linear Models Predicting Teachers' Emphasis on Conceptual Objectives

Variable	Model 1	Model 2	Model 3
Level 1 (teacher)			
Intercept	44.28 (2.03)***	45.08 (1.94)***	45.21 (1.99)***
Sex (female = 1)	2.78 (0.81)***	2.73 (0.80)***	2.73 (0.80)***
Race (reference = White)			
Black	4.96 (1.57)**	4.57 (1.60)**	4.75 (1.56)**
Hispanic	5.41 (1.76)**	4.82 (1.77)**	4.77 (1.71)**
Asian	0.80 (2.46)	–0.31 (2.26)	–0.18 (2.23)
Other	3.78 (1.66)*	3.47 (1.85)†	3.88 (2.05)†
Years of experience (reference = 6–10 years)			
2 years or less	–1.16 (1.16)	–1.26 (1.12)	–1.59 (1.14)
3–5 years	–2.38 (1.18)*	–2.24 (1.14)*	–2.47 (1.14)*
11 years or more	–0.99 (0.95)	–0.88 (0.91)	–0.90 (0.91)
Teaching certificate (reference = no state certification)			
Full certification	2.92 (1.78)	2.56 (1.78)	2.53 (1.78)
Temporary certification	1.65 (2.00)	1.21 (2.00)	1.26 (2.02)
Science certification (reference = no certificate or not in science)			
Science certification	1.60 (1.30)	1.29 (1.26)	1.09 (1.25)
Science certification unknown	1.69 (1.28)	1.85 (1.27)	1.74 (1.27)
Science-related major (reference = no science or science education)			
Science major graduate	4.47 (1.16)***	3.72 (1.18)**	
Science major undergraduate	1.93 (1.02)†	0.97 (0.99)	
Science minor	2.76 (1.56)†	2.74 (1.58)†	
Science education (major or minor)	1.82 (1.67)	1.72 (1.56)	
Other science (science degree but level unclear)	4.16 (2.13)†	2.93 (1.98)	
Any science degree			1.77 (0.90)*
Any education degree	–1.30 (1.11)*	–1.20 (1.04)	–0.99 (1.06)
Professional development workshop		0.11 (0.03)***	0.15 (0.04)***
Interaction with "any science degree"			–0.07 (0.06)
Number of college or university courses		0.75 (0.26)**	0.71 (0.26)**
Level 2 (school)			
School type (reference = public)			
Magnet	–0.67 (1.76)	–0.61 (1.60)	–0.54 (1.56)
Special education	1.45 (2.52)	2.12 (2.44)	2.20 (2.45)
Private	0.32 (1.10)	0.98 (1.09)	1.12 (1.09)
Percentage free and reduced-price lunch	–0.01 (0.02)	–0.01 (0.02)	–0.01 (0.02)
Variance component			
Level 1 variance	71.2	67.6	68.1
Level 2 variance	11.1***	11.7***	11.6***
df	588	588	588
χ^2	778.0	795.2	790.1
Deviance	7749.7	7706.0	7711.5
Number of estimated parameters	25	27	24

Note. Sample size = 1,073 teachers. Unstandardized coefficients are shown with robust standard errors in parentheses. Percentage free and reduced-price lunch, professional development workshop, and number of college or university courses are grand mean centered.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

weaker content knowledge (a remediation effect).

Figure 2 shows how predicted values of teachers' frequency of the use of procedural activities, reporting and writing activities, and hands-on activities in class and the amount of emphasis they gave to conceptual objectives (the *y*-axis) varied by hours of participation in professional development workshops or seminars in science or science education (the *x*-axis) and whether the teachers had degrees in science or science education (the darker line indicates a degree in science or science education; the lighter line indicates a degree in a field outside of science). Although the interaction was negative for the use of reporting and writing activities (β = –.10, $p = .113$; Table 5), the use of hands-on activities (β = –.12, $p = .033$; Table 6), and emphasis on conceptual objectives (β = –.07, $p = .260$; Table 7), it was significant for hands-on activities only at the $p < .05$ level. These patterns can be seen in the convergence in predicted values of instruction for teachers with and without science degrees as their participation hours in science and science education workshops and seminars increase in the top right and bottom two graphs in Figure 2. The use of reporting and writing activities and hands-on activities in class, as well as the amount of emphasis that teachers gave to conceptual objectives for teachers with and without a science degree, converged at between 30 and 40 hours of content-related professional development. Thus, participation in a substantial amount of professional development focused on science or science education is associated with increased use of reform-oriented instructional practices among science teachers of eighth grade students, with an even stronger effect on use of hands-on activities among teachers who did not have a strong science focus as part of their formal education.

Finally, teachers' instructional practices and goals were broadly similar across our teacher demographic and school control variables, although there were a few notable patterns. First, eighth grade science teachers in magnet schools were less likely to report using procedural activities than their counterparts in regular public schools (β = –6.70, $p = .001$; Model 1, Table 4), whereas the reports of using reporting and writing activities and hands-on activities, as well as emphasizing conceptual goals, were not statistically significant between these two types of schools. It is notable that eighth grade science teachers in private schools reported lower use of hands-on activities (β = –2.89, $p = .006$; Model 1, Table 6) and increased use of reporting and writing activities (β = 1.95, $p = .081$; Model 1, Table 5) than their counterparts in regular public schools; schools with greater percentages of students receiving free or reduced-price lunches reported using hands-on activities less frequently as well (β = –.06, $p = .002$; Model 1, Table 6), suggesting that access to laboratory resources may affect the kinds of science instruction taught in high-poverty schools.

Female teachers tended to emphasize reform-oriented activities more than male teachers. Black teachers also emphasized procedural activities more than White teachers, although they were also more likely to emphasize conceptual goals and report more frequent use of reporting and writing activities. Patterns of response for Hispanic teachers were similar to those for Black teachers for time spent on procedural activities, emphasis on conceptual objectives, and time spent on reporting and writing activities, although only the latter two were statistically significant at the $p < .05$ level. Hispanic teachers were more likely than White teachers to report spending time on hands-on activities.

Limitations of the Analysis

The NAEP achievement data were not used in this study. Although we are sensitive to the criticism that teachers' self-reports of their instruction may be less convincing to policy makers than change in student achievement, we focused on instruction as the dependent variable in this study for two reasons. First, the NAEP data are not appropriate for measuring the relationship between credentials and increases in student achievement. Although within- and between-school variation in NAEP achievement scores can be modeled within a hierarchical linear modeling framework, there is no way to control for the prior achievement of the students, and as a result we would not be able to establish the direction of causality between teachers' use of different types of instruction and their students' NAEP achievement scores. In fact, the literature on tracking strongly suggests that teachers alter their

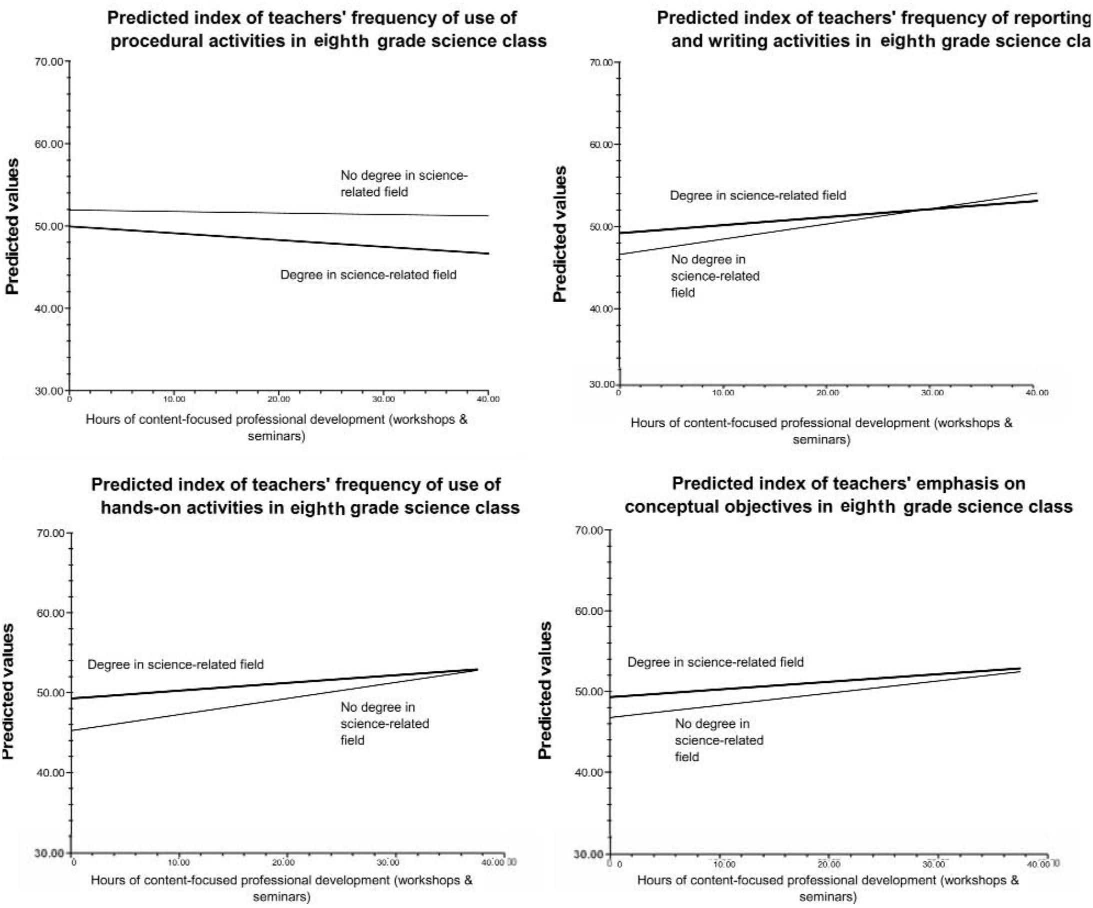

FIGURE 2. *Predicted values of teachers' self-reported frequency of use of different types of instructional activities and level of emphasis on conceptual objectives by hours of content-focused professional development and whether the teacher had a degree in a science-related field.*

instruction on the basis of the ability levels of students in their classes, and we could not adequately separate whether the choice of instructional practices was influenced by initial achievement of students in the class or associated with the change in student achievement during the school year.

Second, we believe that reform-oriented teaching is a relevant policy variable of interest to the education policy community. The "highly qualified" provisions of NCLB are predicated on the idea that better qualifications result in "better" teaching, and that in turn results in increased student achievement. We purposely focused on the link between credentials and instruction to take advantage of a national data set that allows insight into how well credentials are indicators for reform-oriented teaching and how that relationship might work (e.g., the mediating role of professional development). We view these links as policy relevant, in that they are fundamental assumptions of NCLB. We hope that providing some insight into how they work will be useful and informative for those designing policy with credentials as the cornerstone of teacher qualifications requirements.

Discussion

Although inquiry has long been used to characterize good science teaching and learning (Anderson, 2002), few quantitative studies have examined the "technical" barriers preventing the widespread use of inquiry teaching (Anderson, 1996). The technical dimension includes "limited ability to teach constructively, prior commitments (e.g., to a textbook), the challenges of assessment, difficulties of group work, the challenges of new teacher roles, the challenges of new student roles, and inadequate in-service education" (Anderson, 2002, p. 9). Although we measured only a limited range of activities and objectives that eighth grade teachers self-reported using in science classes, we found relatively strong associations between reform-oriented practice and the majors and degrees that teachers earned as part of their formal schooling, as well as their current levels of participation in content-oriented professional development activities. This finding is consistent with earlier findings by Druva and Anderson (1983) that student outcomes in science are positively associated with teacher preparation in education and academic work, particularly in science training, and that the relationship between teachers' training in science and cognitive student outcomes is progressively higher when they have taken higher level science courses.

Furthermore, we found that teachers without degrees in science or science education had similar use rates of hands-on activities in their science classes if they reported participating in sustained (over 35 hours) science or science-focused professional development. These findings should be encouraging for policy makers who want to increase the level of reform-oriented teaching in science classrooms, because the two main policy levers currently being used to improve teaching quality—requiring a degree in the subject taught and encouraging teachers who are not "highly qualified" to take content-focused professional development—are associated with increased use of the kinds of science teaching advocated by the NSES. This finding is notable, because we cannot distinguish in the NAEP data what specific "science" content was covered in these courses. It is likely that the association would be even stronger if the professional development focuses on learning to teach an inquiry-oriented curriculum (Anderson, 2002), emphasizes active learning (Garet et al., 2001; Putnam & Borko, 1997), is grounded in teachers' own instructional practices and curriculum, is reinforced in their classrooms, and is supported by evaluative feedback (Corcoran, McVay, & Riordan, 2003). An additional aspect of the quality of professional development that we were unable to address is the degree to which the curricular materials themselves can be designed to promote teacher learning in addition to student learning, what Davis and Krajcik (2005) called educative curriculum materials. These types of materials can help better scaffold the content that science teachers encounter in professional development and integrate it more effectively into teachers' own instructional practices.

Although our results support the promise of professional development for influencing teachers' use of inquiry in their teaching, they also suggest that one-shot workshops on teaching content are not likely to have large impacts on increased use of reform-oriented strategies. Closing the gaps in emphasis on the use of reporting and writing activities, the use of hands-on activities, and emphasis on conceptual objectives between teachers with degrees in science and those without degrees in science could take as much as 40 hours of participation in content-focused workshops or seminars. This is consistent with Corcoran et al.'s (2003) finding that K–8 teachers need at least 100 hours of professional development to be competent and comfortable delivering kit-based inquiry programs. Districts and schools that provide incentives for teachers to participate in content-focused activities and that focus their professional development programs on content-based activities have the potential to increase science teachers' emphasis on inquiry teaching, helping to close these gaps.

The relationship between holding regular teacher certification and inquiry in teaching is weaker and statistically significant only for the use of hands-on activities. Although full certification is one of the most clearly prescribed requirements for "highly qualified" status, our study suggests that it may not be a sufficient proxy for teachers' use of reform-oriented teaching strategies. Furthermore, although state-level certification to teach science was an indicator of teachers' use of reporting and writing activities, the use of hands-on activities, and emphasis of conceptual objectives (results not shown), the sizes of these effects are reduced, and some of them are no longer statistically significant when degree field and participation in professional development are controlled for. These analyses suggest that certification alone is not the best proxy for preparedness for whether a teacher is going to implement reform-oriented teaching, which is consistent with similar findings in math (Smith et al., 2005).

How best to reform science instruction has been debated by scientists and science educators for decades, although how science is taught is likely to garner wider public policy interest as science achievement becomes part of the formula for calculating whether schools and districts are making adequate yearly progress. Although a focus on inquiry in science is not mentioned in NCLB, both the requirement that states adopt assessments aligned with standards and the fact that state standards are often based on NSES imply an implicit foundation for inquiry to be at the center of reform-oriented teaching. Although we acknowledge that inquiry-oriented instruction is only one of several teacher competencies, such as strong content knowledge or effective classroom management skills, how teachers' knowledge and skills relate to their teaching is an important policy question. The results from our study are consistent with the main findings of the recent National Research Council study *Taking Science to School: Learning and Teaching Science in Grades K-8* (Duschl, Schweingruber, & Shouse, 2007): (a) "student learning of science depends on teachers having adequate knowledge of science," (b) "in order for K-8 teachers to teach science as practice they will need sustained science-specific professional development in preparation and while in service," and (c) "achieving science proficiency for all students will require a coherent system that aligns standards, curriculum, instruction, assessment, teacher preparation, and professional development for teachers across the K-8 years" (p. 296). Although the direction of causality between participation in professional development and teachers' use of inquiry-oriented instruction cannot be ascertained by the cross-sectional survey data analyzed here, it is clear that this is an area ripe for research using more experimental methods and the investment of resources to do a randomized field trial of sufficient scale to inform policy making.

Notes

[1]Small schools, nonpublic schools, and those with high concentrations of Black and/or Hispanic students were oversampled in National Assessment of Educational Progress (NAEP) (U.S. Department of Education, 2001).

[2]If our analyses were at the student level (e.g., science achievement was the dependent variable), we could have nested students within teachers and used student weights for the analysis. Because eighth grade students in the same science class were linked to the same teacher, the student was not an appropriate unit of analysis for our study, in that there would be no variability in instructional practices of science teachers (our dependent variable) among students in the same science class.

[3]We excluded four questions that represented teaching strategies that could not be assigned as either procedural or conceptual: "How often do you use computers for instruction in science?" "How much emphasis do you give to using technology as a scientific tool?" "When you teach, how often do you use computers for science?" and "When you teach, how often do you show a science videotape of a science television program?" To verify that exclusion was reasonable, we conducted a confirmatory factor analysis with and without the four items. Each item loaded on multiple factors, and when they were excluded, the four-factor solution of frequency of the use of procedural activities, reporting and writing activities, and hands-on activities in class and the amount of emphasis that teachers give to conceptual objectives was confirmed.

[4]Probationary status usually represents new teachers who are currently serving their probation periods before full certification.

[5]The question on the NAEP teacher survey that asked teachers to indicate their years of experience teaching science provided response categories that divided years of experience into blocks rather than asking teachers to indicate the exact number of years they had taught science.

References

Abd-el-khalick, F., Bell, R., & Lederman, N. (1998). *The nature of science and instructional practice: Making the unnatural natural*. New York: John Wiley.

Anderson, R. D. (1996). *Study of curriculum reform*. Washington, DC: U.S. Government Printing Office.

Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education*, *13*(1), 1–12.

Anderson, R. D., & Helms, J. F. (2001). The ideal of standards and the reality of schools: Needed research. *Journal of Research in Science Teaching*, *38*(1), 3–16.

Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly*, *58*(2), 210–240.

Aquilino, W. S. (1998). Effects of interview mode on measuring depression in younger adults. *Journal of Official Statistics*, *14*(1), 15–29.

Aquilino, W., & LoSciuto, L. (1990). Effects of interview mode on self-reported drug use. *Public Opinion Quarterly*, *54*, 362–395.

Ball, D. L. (1990). The mathematical understandings that prospective teacher bring to teacher education. *Elementary School Journal*, *90*(4), 449–466.

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the National Science Foundation's Local Systemic Change Through Teacher Enhancement initiative. *Journal of Research on Science Teaching*, *44*(3), 375–395.

Blumfeld, P. C., Krajcik, J. S., Marx, R. W., & Soloway, E. (1994). Lessons learned: How collaboration helped middle grade science teachers learn project-based instruction. *Elementary School Journal*, *94*(5), 539–551.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229–270). Cambridge, MA: MIT Press.

Burstein, L., McDonnell, L. M., Van Winkle, J., Ormseth, T., Mirocha, J., & Guitton, G. (1995). *Validating national curriculum indicators*. Santa Monica, CA: RAND.

Carlsen, K. (1990). Teachers' knowledge and learning to teach. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 291–310). New York: Macmillan.

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

Cavanagh, S. (2005, March 30). As test date looms, educators renewing emphasis on science. *Education Week*, *24*(29), 14–15.

Cochran, K., & Jones, L. (1998). The subject matter knowledge of preservice science teachers. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 707–718). Dordrecht, the Netherlands: Kluwer Academic.

Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist*, *27*, 291–315.

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, D. K., & Hill, H. C. (2001). *Learning policy: When state education reform works*. New Haven, CT: Yale University Press.

Cohen, D. K., & Spillane, J. P. (1993). Policy and practice: The relationship between governance and instruction. In S. H. Fuhrman (Ed.), *Designing coherent education policy: Improving the system* (pp. 35–95). San Francisco: Jossey-Bass.

Corcoran, T. B. (1995). *Transforming professional development for teachers: A guide for state policymakers*. Washington, DC: National Governors' Association.

Corcoran, T., McVay, S., & Riordan, K. (2003, December). *Getting it right: The MISE approach to professional development* (CPRE Research Report Series RR-055). Philadelphia: University of Pennsylvania, Graduate School of Education, Consortium for Policy Research in Education.

Corcoran, T., Shields, P., & Zucker, A. (1998, March). *The SSIs and professional development for teachers: Evaluation of NSF's Statewide Systemic Initiatives (SSI) Program*. Arlington, VA: National Science Foundation, Division of Research, Evaluation, and Communication, Directorate for Education and Human Resources.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, *8*(1). Available at http://olam.ed.asu.edu/epaa/v8n1/

Davis, E. A., & Krajcik, J. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 3–14.

Desimone, L., Porter, A. C., Garet, M., Suk Yoon, K., & Birman, B. (2002). Does professional development change teachers' instruction? Results from a three-year study. *Educational Evaluation and Policy Analysis*, *24*(2), 81–112.

Desimone, L. M., Smith, T. S., Hayes, S., & Frisvold, D. (2005). Beyond accountability and average math scores: Relating multiple state education policy attributes to changes in student achievement in procedural knowledge, conceptual understanding and problem solving in mathematics. *Educational Measurement: Issues and Practice, 24*(4), 5-18.

Dillman, D. A., & Tarnai, J. (1991). Mode effects of cognitively designed recall questions: A comparison of answers to telephone and mail surveys. In P. N. Beimer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 367–393). New York: John Wiley.

Druva, C., & Anderson, R. (1983). Science teacher characteristics by teacher behavior and by student outcome: A meta-analysis of research. *Journal of Research in Science Teaching*, *20*(5), 467–479.

Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment*, *4*(1), 37–73.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in Grades K-8*. Washington, DC: National Academies Press.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. A., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education*, *27*(4), 403–434.

Fowler, F. J., Jr., Roman, A. M., & Di, Z. X. (1998). Mode effects in a survey of Medicare prostate surgery patients. *Public Opinion Quarterly*, *62*, 976–989.

Garet, M., Birman, B., Porter, A. C., Yoon, K., & Desimone, L. (2001). What makes professional development effective? Analysis of a national sample of teachers. *American Education Research Journal*, *38*(3), 915–945.

Gitomer, D. H., & Duschl, R. A. (1995). Moving toward a portfolio culture in science education. In M. S. Glynn & R. Duit (Eds.), *Learning science in the schools: Research reforming practice* (pp. 299–325). Mahwah, NJ: Lawrence Erlbaum.

Glasson, G. (1989). The effects of hands-on and teacher demonstration laboratory methods on science achievement in relation to reasoning ability and prior knowledge. *Journal of Research in Science Teaching*, *26*(2), 121–131.

Hashweh, M. Z. (1987). Effects of subject matter knowledge in the teaching of biology and physics. *Teaching and Teacher Education*, *3*, 109–120.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., et al. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher*, *25*(4), 12–21.

Hochstim, J. R. (1967). A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, *62*, 976–987.

Hollweg, K. S., & Hill, D. (Eds.). (2003). *What is the influence of the national science education standards? Reviewing the evidence: A workshop summary*. Washington, DC: National Academies Press.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667.

Lott, G. W. (1983). The effect of inquiry teaching and advanced organizers upon student outcomes in science education. *Journal of Research in Science Teaching*, *29*(5), 437–451.

Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin.

Luft, J. (2001). Changing inquiry practices and beliefs: The impact of an inquiry-based professional development program on beginning and experienced secondary science teachers. The effects of professional development on science. *International Journal of Science Education*, *23*(5), 517–534.

Luft, J., & Pizzini, E. (1998). *The demonstration classroom in-service: Changes in the classroom*. New York: John Wiley.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum.

Marx, R. W., & Harris, C. J. (2006). No Child Left Behind and science education: Opportunities, challenges, and risks. *Elementary School Journal*, *106*(5), 467–477.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, *21*(1), 29–45.

Mayer, D. P., Mullens, J. E., & Moore, M. T. (2001). *Monitoring school quality: An indicators report* (NCES 2001-30). Washington, DC: U.S. Department of Education, U.S. Government Printing Office.

Mechling, K. R., & Oliver, D. L. (1983). *Promoting science among elementary school principals*. Washington, DC: National Science Teachers Association.

Mullens, J. (1995). *Classroom instructional processes: A review of existing measurement approaches and their applicability for the teacher follow-up survey* (NCES 95-15). Washington, DC: National Center for Education Statistics.

Mullens, J. E., & Gayler, K. (1999). *Measuring classroom instructional processes: Using survey and case study field test results to improve item construction* (NCES 1999-08). Washington, DC: National Center for Education Statistics.

Mullens, J., & Kasprzyk, D. (1996). Using qualitative methods to validate quantitative survey instruments. In *1996 proceedings of the Section on Survey Research Methods* (pp. 638–643). Alexandria, VA: American Statistical Association.

Mullens, J., & Kasprzyk, D. (1999). *Validating item responses on self-report teacher surveys*. Washington, DC: U.S. Department of Education.

Murnane, R. J., & Phillips, B. R. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. *Economics of Education Review*, *1*, 691–693.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press. Retrieved July 16, 2007, from http://www.nap.edu/readingroom/books/nses/html/1.html.

National Research Council. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press.

Newmann, F., Smith, B., Allensworth, E., & Bryk, A. (2001, January). *School instructional program coherence: Benefits and challenges*. Chicago: Consortium on Chicago School Research.

No Child Left Behind Act of 2001, Pub. L. 107-110 (2001).

Putnam, R. T., & Borko, H. (1997). Teacher learning: Implications of new views of cognition. In B. J. Biddle, T. L. Good, & I. F. Goodson (Eds.), *The international handbook of teachers and teaching* (Vol. 2, pp. 1223–1296). Dordrecht, Netherlands: Kluwer.

Raudenbush, S. W., Rowan, B., & Cheong, Y. (1993). Higher order instructional goals in secondary schools: Class, teacher, and school influences. *American Educational Research Journal*, *30*, 523–553.

Remillard, J. T. (2000). Can curriculum materials support teachers' learning? Two fourth-grade teachers' use of a new mathematics text. *Elementary School Journal*, *100*(4), 331–350.

Reynolds, A. (1995). The knowledge base for beginning teachers: Education professionals' expectations versus research findings on learning to teach. *Elementary School Journal*, *95*(3), 199–221.

Rhine, S. (1998). The role of research and teachers' knowledge base in professional development. *Educational Researcher*, *27*(5), 27–31.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (1998). *Teachers, schools and academic achievement*. Paper presented at the meeting of the Association for Public Policy Analysis and Management, New York.

Rosebery, A. S., Warren, B., & Conant, F. R. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *Journal of the Learning Sciences*, *2*(1), 61–94.

Rowan, B., Chiang, F., & Miller, R. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, *70*(4), 256–284.

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

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *A splintered vision: An investigation of U.S. science and mathematics education*. Boston: Kluwer Academic.

Schneider, R. M., & Krajcik, J. (2002). Supporting science teacher learning: The role of educative curriculum materials. *Journal of Science Teacher Education*, *13*(3), 221–245.

Shavelson, R. J., Webb, N. M., & Burstein, L. (1986). Measurement of teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 50-91). New York: MacMillan.

Shymansky, J. A., Kyle, W. C., & Alport, J. M. (1983). The effects of new science curricula on student performance. *Journal of Research in Science Teaching*, *20*(5), 387–404.

Simon, M. A., & Schifter, D. (1991). Towards a constructivist perspective: An intervention study of mathematics teacher development. *Educational Studies in Mathematics*, *22*, 309–331.

Smerdon, B., Burkam, D., & Lee, V. (1999). Access to constructivist and didactic teaching: Who gets it? Where is it practiced? *Teachers College Record*, *101*(1), 5–34.

Smith, T., Desimone, L., & Ueno, K. (2005). "Highly qualified" to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school math. *Educational Evaluation and Policy Analysis*, *27*(1), 75–109.

Smithson, J. L., & Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum—The Reform Up Close study* (CPRE Research Report Series #31). Madison: University of Wisconsin, Consortium for Policy Research in Education.

Supovitz, J., & Turner, H. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching*, *37*(9), 963–980.

Sykes, G. (1996). Reform of and as professional development. *Phi Delta Kappan*, *77*(7), 465–489.

Tal, T., Krajcik, J. S., & Blumenfeld, P. (2006). Urban schools teachers enacting project-based science. *Journal of Research in Science Teaching*, *43*(7), 722–745.

Tobin, K., & Tippins, D. (1993). Constructivism as a referent for teaching and learning. In K. Tobin (Ed.), *The practice of constructivism in science education* (pp. 3-21). Hillsdale, NJ: Lawrence Earlbaum & Associates.

U.S. Department of Education. (1999). *The NAEP 1996 technical report* (NCES 199-452). Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. (2000). *The nation's report card: Mathematics 2000* (NCES 201-517). Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. (2003). No Child Left Behind: A toolkit for teachers. Retrieved January 11, 2004, from http://www2.edtrust.org/NR/rdonlyres/C638111D-04E3-4C0D-9F68-20E7009498A6/0/tellingthetruthteachers.pdf

U.S. Department of Education. (2007a). *Fiscal year 2001-2008 state tables for the U.S. Department of Education*. Available at http://www.ed.gov/about/overview/budget/statetables/index.html

U.S. Department of Education. (2007b). *Science framework for the 2009 national assessment of* *educational progress*. Washington, DC: U.S. Government Printing Office.

Von Secker, C. (2002). Effects of inquiry-based teacher practices on science excellence and equity. *Journal of Educational Research*, *95*(3), 151–160.

Von Secker, C., & Lissitz, R. (1999). Estimating the impact of instructional practices on student achievement in science. *Journal of Research in Science Teaching*, *36*(10), 1110–1126.

Wise, K. C., & Okey, J. R. (1983). A meta-analysis of the effects of various science teaching strategies on achievement. *Journal of Research in Science Teaching*, *20*(5), 419–435.

Wright, J. C., & Wright, C. S. (1998). A commentary on the profound changes envisioned by the National Science Education Standards. *Teachers College Record*, *100*(1), 122–143.

Authors

THOMAS M. SMITH is an assistant professor of Public Policy and Education in the Department of Leadership, Policy, and Organizations, # 514 Peabody College,Vanderbilt University, Nashville TN 37203-5721; tom.smith@vanderbilt.edu. His research focuses on the organizational and policy contexts of teaching and teacher quality.

LAURA M. DESIMONE is currently an associate professor at the Graduate School of Education at the University of Pennsylvania., 3700 Walnut Street, Philadelphia, PA 19104-6216; lauramd@gse.upenn.edu. Her research focuses on policy effects on teaching and learning, policy implementation, and improving surveys.

TIMOTHY L. ZEIDNER is currently a research associate for the Tennessee State Board of Education, 9th Floor Andrew Johnson Tower, 710 James Robertson Parkway, Nashville, TN 37243; Tim.Zeidner@state.tn.us. His research interests focus on varying educational policies that influence the likelihood of students accessing and succeeding in postsecondary educational institutions, from teacher certification to school choice measures.

ALFRED C. DUNN is currently a third year doctoral student in the Department of Education Leadership and Policy and member of the Experimental Education Research Training Program (ExPert), which is supported by the Institute for Education Sciences (IES), at Vanderbilt University, Peabody #329, 230 Appleton Place, Nashville, TN 37203; alfred.c.dunn@vanderbilt.edu. His research focus is on secondary mathematics and science teachers, teacher quality, schools as organizations, teacher job satisfaction, and teacher turnover.

MONICA BHATT is a Master's of Public Policy student in the Department of Leadership, Policy, and Organizations at Vanderbilt University, Peabody College, Peabody #329, 230 Appleton Place, Nashville, TN 37203; m.bhatt@vanderbilt.edu. As a graduate student she has been involved in projects focusing on issues of teacher quality, teacher induction, and mentoring.

NATALIYA L. RUMYANTSEVA is currently a doctoral candidate in the Department of Leadership, Policy and Organizations at Vanderbilt University, Peabody College, Peabody #329, 230 Appleton Place, Nashville, TN 37203; n.rumyantseva@vanderbilt.edu. She is part of several research projects that span a wide variety of interests that include organizational sociology and organizational context of teaching and learning.

Manuscript received August 15, 2006
Final revision received June 20, 2007
Accepted June 21, 2007