Promoting Inquiry-Based Instructional Practice: The Longitudinal Impact of Professional Development in the Context of Systemic Reform

JONATHAN A. SUPOVITZ,
DANIEL P. MAYER, and JANE B. KAHLE

Over the past decade, national and state policy makers have wagered heavily on systemically reforming science and mathematics education through inquiry-based instructional approaches. This study examines the impact in one state of intensive science and mathematics professional development for teachers in the context of statewide systemic reform. Based on a unique longitudinal data set, the study models the impact of professional development on teachers' attitudes toward inquiry-based instruction, their capacity to adopt inquiry-based teaching strategies, and their classroom use of inquiry-based instructional practices. The findings are remarkably consistent across both subjects. Teachers' attitudes, preparation, and practices all showed strong, positive, and significant growth from preprofessional development to the following spring. Furthermore, these gains were sustained over several years following their involvement. These findings provide a promising indication that large-scale, high-quality, intensive training set within a context of standards-based systemic reform can be a powerful mechanism for sustained impact on teachers.

AUTHORS' NOTE: The preparation of this article was funded in part by a grant from the National Science Foundation (NSF), Grant #REC 9602137. The opinions expressed are those of the authors and do not necessarily reflect the position of NSF.

EDUCATIONAL POLICY, Vol. 14 No. 3, July 2000 331-356

THE PROMISE of inquiry-based instruction to change the heart of educational practice has pulsed through at least three major educational reform movements in the United States in the 20th century. On the first two occasions, the Progressive movement of the 1920s and the curriculum reform movement of the 1950s, the reforms failed to take hold because they did not have widespread influence on teachers—the key players in any instructional reform (Elmore, 1996). The third occasion, the systemic reform movement that began in the early 1990s, promoted the use of inquiry-based instruction through sustained professional development efforts. This article examines whether these professional development efforts, embedded in systemic reform, can promote teacher use of inquiry-based instruction.

Inquiry-based instruction is a student-centered pedagogy that uses purposeful extended investigations set in the context of real-life problems as both a means for increasing student capacities and as a feedback loop for increasing teachers' insights into student thought processes. First widely promulgated by John Dewey (1915), the inquiry philosophy was a central tenet of the Progressive reform movement (Cremin, 1961). Progressive education reformers undertook massive curriculum reforms in the 1920s and 1930s in many U.S. school districts but, according to educational historian Larry Cuban (1984), the new curricula "seldom appeared in more than one-fourth of the classrooms in any district that systematically tried to install these various elements" (p. 135). With a few notable exceptions, reformers' efforts of that era had little impact on classroom teaching practices.

The second inquiry-based instructional reform movement occurred in the late 1950s and 1960s. In part spurred by Russia's launching of Sputnik in 1957, the central idea of the reforms was that students should engage in investigations similar to those of professional scientists and discover both the subject content and methods, as well as the processes of constructing their knowledge (Dow, 1991). The reform focused on curriculum codeveloped by university faculty content experts, science educators, and school teachers. Curriculum projects included the Physical Sciences Study Committee's (PSSC) high school physics curriculum, begun in 1956; the Biological Sciences Curriculum Study (BSCS), begun in 1958; and Man: A Course of Study (MACOS), begun in 1959. These curriculum projects received substantial funding from both private foundations and the National Science Foundation (NSF) (Elmore, 1993; Grobman, 1969). But, like the progressive efforts, what resulted had very little impact on the core of U.S. schooling (Elmore, 1996).

The third major inquiry-based reform effort of the 20th century has been underway since the early 1990s, when reformers began advocating inquiry-based instruction as a primary element of systemic reform. Systemic reformers seek to develop coherent educational policies through coordinated governance structures that align the various components of the educational system (i.e., standards, assessments, curricula, professional development, etc.) to provide reinforcing and complementing policies (Vinovskis, 1996). Driven by the release of the mathematics standards (National Council of Teachers of Mathematics, 1989, 1991) and the science standards (National Research Council, 1996), inquiry-based instruction is a central tenet of systemic reform.

A principal example of the pivotal role of inquiry-based teaching in systemic reform is the NSF's statewide systemic initiative (SSI) program, which provided grants of approximately $10 million over 5 years to 24 states and Puerto Rico to reform their educational systems. Perhaps learning from the mistakes of the 1950s curriculum-centered reforms, which popularized the term *teacher-proof curriculum* (Cremin, 1961), the SSI reform movement is driven by professional development that, in the best of cases, both teaches and models inquiry-based instruction. An SSI evaluation that looked across sites found that "increasing the opportunities for professional development of teachers and enhancing their quality were central strategies for almost all of the SSIs" (Corcoran, Shields, & Zucker, 1998, p. 3). In fact, professional development represented the single largest category of expenditures across the SSIs (Marder, 1996). In 1994 alone, SSIs reported providing professional development to nearly 50,000 teachers, approximately 8% of the public school teachers in the SSI states (Zucker, Shields, Adelman, & Powell, 1995).

This article explores the potential fate of the third movement by assessing the impact of inquiry-based professional development in Ohio, one of the first cohort SSIs, funded in 1991. Ohio's systemic reform strategy centered on delivering 6-week inquiry-based professional development institutes to science and mathematics teachers via a regional infrastructure. Intensive content institutes in physical science, mathematics, and life science were followed by six seminars spread throughout the academic year that focused on grade-appropriate curriculum, equity issues, and authentic assessment strategies. As Ohio's SSI matured, other professional development programs were implemented to reach more teachers. However, the 6-week institute remained the SSI's core professional development strategy.

Using Ohio's experience as an example, our analysis examines the following questions: What are both the short- and long-term impacts of systemic reforms that center their efforts on inquiry-based professional development? More specifically, what are the reform's impacts on teachers' attitudes toward inquiry-based instruction, their preparation to adopt inquiry teaching techniques, and their classroom use of inquiry instructional practices? Furthermore, what are the implications of these policies for educational reformers? To answer these questions, we employed a hierarchical linear growth model that allowed us to examine both the short- and long-term impact of Ohio's professional development efforts.

First, we describe the context, methods, and findings in greater detail. Second, we detail Ohio's systemic reform initiative. Third, we describe the methods used in this study, including a description of the measures, a portrait of the sample, an account of the analytic approach, and a technical explanation of the construction of the models used for this analysis. Fourth, we reveal the results of the study. Finally, we discuss the importance and implications of the findings.

OHIO'S PROFESSIONAL DEVELOPMENT IN THE CONTEXT OF SYSTEMIC REFORM

Ohio's systemic reform strategy focuses on improving student learning of science and mathematics by providing teachers with the knowledge, skills, and information needed to change their individual teaching practices; by helping teachers to affect lateral change in their schools and districts as they serve as workshop leaders and in other capacities; and by helping teachers to learn how to influence curricular and assessment policies in their districts and state. Using a regional delivery model through existing higher education institutions, Ohio's SSI, called Discovery, began sponsoring inquiry-based professional development in mathematics, life science, and physical science in 1992. Discovery concentrated initially on reaching middle school teachers and then, over time, increased emphasis on elementary and high school teachers. To date, Discovery has provided professional development to approximately 20% of the middle school teachers in the state.

Discovery was based on two premises. First, that classroom teachers—particularly elementary and middle school teachers—were woefully underprepared to teach science and mathematics effectively. Second, that the needs of an increasingly diverse student population required teachers to employ instructional strategies designed to increase the motivation of that student population. Enhancing teachers' knowledge of science and mathematics and improving their skills in the use of inquiry teaching became the centerpiece of Ohio's Statewide Systemic Initiative. Discovery's strategy was supported by evidence that suggested the effectiveness of inquiry instruction (Arons, 1989; Bybee, 1993; McDermott, 1996). That literature also indicated that if teachers were to effectively use inquiry, then they needed to learn by inquiry (Arons, 1989; McDermott, 1996). Research studies concerning gender and racial/ethnic differences in participation and achievement in science and mathematics also suggested that inquiry motivated underrepresented students (Atwater, 1994; Fennema, 1990; Kahle, 1996).

Briefly, teaching by inquiry seeks to build on a student's natural inquisitiveness to develop the scientific (including mathematical) knowledge and thinking skills that enhance one's daily life and make a career in science, mathematics, or engineering possible. In contrast to traditional instructional methods that emphasize what is known, teaching by inquiry stresses what is unknown, particularly to the student. In inquiry instruction, students and teachers collectively pose a question, collect evidence, and develop one or more solutions. Inquiry instruction can take many forms, varying from structured inquiry in which there is a known outcome to open-ended inquiry, which may have one or more outcomes. That is, questions are posed, and students seek answers though investigations, including experimentation, collecting evidence, and problem solving. Open-ended questioning and cooperative group work are cornerstones of inquiry teaching forming the foundation of the professional development provided by Discovery.

Discovery's professional development institutes consist of an intensive 6-week, 8 graduate semester hours (approximately 160 contact hours), summer professional development experience. The institutes, conducted by academic leadership teams that include higher education faculty and K-12 lead teachers, are based on *Physics by Inquiry* (McDermott, 1996). The model used in the *Physics by Inquiry* curriculum is closer to structured inquiry. The goals of the institutes are to expand the content knowledge of teachers through inquiry-based instruction, model inquiry teaching so teachers can experience how this pedagogical tool is applied to real-world concepts, and relate course content to national science and mathematics standards. Participating teachers received stipends for participation.

During separate 6-week institutes in physics and mathematics, teachers were divided into cooperative learning groups. Although teachers of Grades 5 through 9 were given preference, teachers of all grade levels eventually attended the institutes. Instructors attempted to construct the cooperative groups to include different grade levels. Daily, new problems were posed, generally through the curriculum in physics and by the students and instructors in mathematics. Groups were left to design experiments or try various problem-solving strategies. When a group thought that it had found a solution, it had to explain and justify its thinking to the instructor. Algorithms were not used; rather, the focus was on learning through doing. Teachers kept journals that chronicled their experiences and, in the physics and life science institutes, responded to embedded assessments. As the number of sites offering institutes grew from 2 to 28, quality and uniformity were ensured by requiring all instructors to participate as students in the institute that they were to teach.

Ongoing school district support is an important component of professional development. Following their intensive summer professional development, teachers participated in academic year seminars (2 graduate semester hours of credit), receiving release time from their districts to meet for 6 days throughout school year. These follow-up sessions focused on sharing strategies, adapting curricula, and addressing the issues that teachers face when they try to implement inquiry-based instruction. Discovery reports that attendance at these follow-up sessions was in excess of 90%.

After the summer institutes, Discovery teachers were supported in a variety of ways. First, a regional infrastructure was developed that provided sustained local support. That is, teachers who had gone through the professional development year received on-site visits and on-demand support from regional leadership teams composed of doctoral-level scientists and mathematicians and teacher-leaders. Based on teacher feedback, these support services varied widely in quality across the state. Teachers also had access to their peers through an electronic chat service called Discovery Net, periodic newsletters, and annual 2-day conferences.

Finally, it is important to understand Discovery's role within the context of science and mathematics education reform in Ohio. Discovery built on the strong foundation laid by both the National Council of Teachers of Mathematics (NCTM) (1989) standards and the tradition of mathematics reform in the state, which emphasized the importance of inquiry. For example, a high-stakes state proficiency test in mathematics was in place prior to the SSI, and the state had an approved model curriculum in mathematics. In science, Discovery's leadership played an important role in the development of the science model curriculum and state proficiency tests. In addition, Discovery's emphasis on middle schools resulted in new middle school certification programs in science and mathematics. Furthermore, Ohio is in the process of changing from teacher certification to licensure—a change that recognizes the importance of substantive professional development.

METHODS

Our analysis of the impact of Discovery is based on teacher survey data collected between 1992 and 1995 as part of Horizon Research, Inc.'s (HRI) evaluation of the Ohio SSI. The survey instruments that were used to collect data from Discovery participants were based on those used in HRI's 1993 National Survey of Science and Mathematics Education (Weiss, Matti, & Smith, 1994). The surveys asked teachers questions about their attitudes, beliefs, and inquiry-based teaching practices, as well as for demographic information about both themselves and their schools.

To evaluate whether teacher attitudes toward inquiry-based instruction, preparation to implement inquiry-based instruction, and classroom use of inquiry-based teaching practices changed over time, we developed constructs of the three phenomena. We measured whether they changed over time while controlling for a series of individual and school characteristics.

Although we hypothesize that teachers must have a positive attitude toward inquiry-based instruction before they feel prepared to use inquiry-based instruction, and attitudes and preparation must both precede changes in inquiry-based instructional practices, we examined the three constructs independently. Both individual teacher characteristics (involvement in reform activities, years of experience, gender, ethnicity, grade level) and school characteristics (school climate, percentage of minority students in the school, and whether the school was public or private) were used to predict growth at each time period. In the rest of this section, we discuss issues associated with the reliability and validity of self-reported data. We then describe the sample the variables used in the analysis. Finally, we detail our analytic strategy.

Reliability and Validity of Self-Reported Teacher Survey Data

In the 1980s, there was some question as to the reliability and validity of self-reported teacher survey data. On one hand, surveys were more cost-effective than observations, interviews, artifacts, or teacher logs as a way to collect data on teaching practices. On the other hand, it was unclear whether teachers could report validly on the schooling process or whether curricular practice could be measured accurately on a survey without observing the interactions between teachers and students (Burstein et al., 1995).

Researchers in the 1990s have learned much about the strength of surveys to collect data accurately on teaching practices. Porter, Kirst, Osthoff, Smithson, and Schneider (1993) examined the consistency between survey responses pertaining to instructional style and detailed teacher logs describing actual lessons, and they concluded that substantial overlap existed. Burstein et al. (1995) used interviews, observations, daily teacher logs, and classroom artifacts to validate survey data. They found that survey data could depict topical content and instructional strategies validly but that instructional goals were more difficult to capture accurately through survey research. Mayer (1999) explored the reliability and validity of survey data as part of his study of the relationship between NCTM-based teaching practices and student achievement. Assessing the reliability of his surveys, Mayer surveyed a group of teachers twice in a 4-month period and found a correlation of .69 between responses on the two administrations. To validate his

Table 1
Response Rates of Discovery Science and Mathematics Participants for Each Time Period

Science				*Mathematics*			
Time 1	*Time 2*	*Time 3*	*Time 4*	*Time 1*	*Time 2*	*Time 3*	*Time 4*
100%	74%	44%	38%	100%	62%	43%	41%

instruments, Mayer observed a random sample of nine classes and found a correlation of .85 between observational data and survey responses.

Sample

The sample for our analysis consisted of the 1,475 Ohio teachers who participated in Discovery professional development activities between 1993 and 1995. Participants in Ohio's science and mathematics professional development sessions were surveyed at the beginning of their 6-week summer professional development on their attitudes and beliefs about teaching and their classroom practices. They were surveyed again each spring for up to 3 years.

Thus, we have up to 4 data points for three cohorts of Discovery participants. That is, for the 1993 cohort, we have 4 data points (1993-1996); for the 1994 cohort, we have 3 data points (1994-1996); and for the 1995 cohort, we have 2 data points (1995 and 1996). The overall response rates for each of the four time periods covered in this analysis are shown in Table 1. Time 1 can be considered the entire cohort of attendees because this group completed their survey as a captive audience. The first postexperience survey, Time 2, was administered by mail the following spring, about 10 months after the professional development began. The data for Times 3 and 4 were also collected by mail each subsequent spring.[1]

Attitudes, Preparation, and Use of Inquiry-Based Instruction

We wanted to know whether teachers' attitudes toward inquiry-based instruction, their preparation to use inquiry, and their actual use of inquiry-based instructional practices changed after participating in the Discovery professional development activities. Based on the survey questions, composites were constructed to represent teachers' attitudes, preparation, and use of inquiry-based instruction in both science and mathematics. A complete list of the questions that comprised the constructs in both subjects is shown in the appendix.

The composite of teachers' attitudes toward inquiry-based instruction was based on questions about their beliefs about the teaching of mathematics or science. The reliability of these composites ranged from .54 to .64 in mathematics and from .36 to .51 in science, depending on the cohort and year examined.

The composite of teachers' preparation for inquiry-based instruction was based on a series of questions about their feelings of preparedness to organize and facilitate an inquiry-based classroom using such techniques as cooperative learning, technology, and a variety of forms of assessment. The reliability of the preparation composites are far more solid than those of attitudes, ranging from .84 to .88 in mathematics and from .78 to .87 in science, depending on the cohort and year examined.

The composite of teachers' use of inquiry-based instruction was based on a series of questions about the extent to which the students in the teachers' mathematics/science classes were taking part in inquiry-based activities, such as working in small groups, doing hands-on activities, and working on long-term projects. The reliability of the practice composites were again quite strong, ranging from .72 to .85 in mathematics and from .66 to .86 in science, depending on the cohort and year examined.

Finally, to facilitate the interpretation of the magnitude of growth in the three areas and to minimize the effects of slight differences in the wording of the questions underlying the composites of attitudes, preparation, and practice over time, the three outcome composites were standardized to have a mean of zero and a standard deviation of 1. The standardization was done once across the multiple data points for each individual.

Teacher and School Characteristics

Participants in Ohio's summer institutes were predominantly experienced White female middle school teachers from public schools. Table 2 shows the individual characteristics of the participants and the attributes of the schools they attended.[2] Each of these variables was used in the subsequent analysis.

In the mathematics sample, 76% of the participants were women, whereas in the science sample, 69% were women. More than 90% of the participants were White, with about 8% African American.[3] The racial/ethnic composition of the sample roughly matches the demographic proportions in the state as a whole. Approximately half of the teachers in both the science and mathematics samples taught middle school, here defined as teachers who reported teaching in Grades 6, 7, or 8. Approximately one quarter of the teachers reported teaching elementary school (K-5), and one quarter taught high school (Grades 9-12). On average, both the science and mathematics participants had approximately 11 years of teaching experience. However, there was a wide distribution in experience, with each sample having a standard deviation of about 8 years. Teachers reported experience levels ranging from 1 to 32 years.

Table 2
Individual Characteristics of Teachers Receiving Discovery Professional Development

	*Science (*n *= 701)*	*Mathematics (*n *= 603)*
Female teachers (%)	69	76
Male teachers (%)	31	24
White teachers (%)	93	92
African American teachers (%)	7	8
Elementary teachers (%)	24	26
Middle school teachers (%)	52	55
High school teachers (%)	25	20
Years of experience		
M	10.91	11.46
SD	8.15	8.35
Minimum	1.00	1.00
Maximum	32.00	32.00
Professional involvement		
M	1.45	1.23
SD	1.33	1.28
Minimum	0.00	0.00
Maximum	5.00	5.00

A final participant characteristic used in this study was a teacher's involvement in professional activities before he or she participated in Discovery. This measure included participation in activities such as teacher association meetings and curriculum development committees over the previous 5 years. We chose this indicator as a measure of the extent to which individual teachers were active in the teaching profession prior to their training. The reliability of this five-item construct, computed by using Cronbach's internal consistency measure coefficient alpha, was .75.

We also included several of the teachers' school characteristics in our model because we hypothesized that teachers in different school circumstances might adopt inquiry differently. These characteristics are described in Table 3. Ninety-three percent of the teachers in Discovery were from public schools. On average, the schools these teachers taught in were composed of approximately 20% minority students, including African American, Hispanic, and Asian students. There was also a broad variation in the percentage of minority students in the schools of Discovery participants, with responses ranging from no minority students to 100% minority students.

Finally, we included a measure of the climate in which teachers were instructing. This measure assesses the degree to which teachers felt empowered within their school. The questions used to create this variable asked

Table 3
School Characteristics of Teachers Receiving Discovery Professional Development

	*Science (*n *= 701)*	*Mathematics (*n *= 603)*
Public school teachers (%)	93	93
Private school teachers (%)	7	7
Minority students in school (%)		
M	20	21
SD	27	28
Minimum	0	0
Maximum	100	100
School climate		
M	3.15	2.96
SD	0.63	0.60
Minimum	1.00	1.00
Maximum	4.00	5.00

teachers about the extent to which teachers in their particular school contributed to school curriculum decisions; shared ideas, materials, and instructional strategies; and felt supported by their principal. We believe that teachers in more empowered environments would be more likely to adopt inquiry-based instructional practices. The Cronbach's alpha reliability of this construct was .70.

Analytic Strategy

To assess the impact of Discovery, we investigated whether individual teacher's attitudes, preparation, and practice changed over a 4-year time span using hierarchical linear modeling (HLM). Most professional development impact studies, even if they are longitudinal in design, only examine outcomes over a few months. In addition, most longitudinal studies employ only a pretest/posttest design. The limitations of measuring individual change using this approach have been well documented (Willett, 1994). First, a pretest/posttest design does what it is supposed to do poorly. A preworkshop assessment can only control for initial differences among the teachers imperfectly, thereby leading to biased parameter estimates (Rogosa, Brandt, & Zimowski, 1982). An additional source of bias derives from the correlation between the pre- workshop score and any unobserved influences on teaching attitudes and behavior (Willett, 1994), such as the teacher's instructional level (i.e., elementary, middle, or high school), gender, or prior involvement in other professional development activities. Thus, by only controlling for differences with a preworkshop assessment, the measure of teacher development probably reflects more teacher background characteristics than changes in attitudes and teaching style.

To avoid these problems, we looked at individual growth over time by gathering as many as four measures of attitude, preparation, and practice. Having multiple measures of each construct allowed us to estimate individual growth trajectories by using growth curve modeling. Growth curve modeling is one of the most powerful methods for studying individual change because it does not suffer from the limitations of the pre/post design, and it allows for the estimation of variation in teachers' starting points and growth in attitudes, preparation, and practices while controlling for important demographic differences among the teachers (Rogosa & Saner, 1995). Additionally, a growth model has the advantage of being able to predict individual growth rates as a function of personal and environmental characteristics, in which each characteristic may predict growth differently depending on the period of time. For example, teachers with more experience may have a more advanced initial starting point, but their experience may have no relationship to the rate at which they grow.

The models we fit predicted measures of teacher attitudes, preparation, and practice as a function of both time and teacher and school characteristics to test whether the trajectories vary across teachers. For this analysis, we used a two-level hierarchical model using the HLM/3L software of Bryk, Raudenbush, and Congdon (1996). The HLM/3L software can include predictor variables associated with individual teachers and schools, and it can incorporate variation occurring at more than one period of time that is distinct from measurement error variance.

Growth rates were not anticipated to be constant between the preworkshop survey and the fourth administration of the survey. The reform indicators were expected to grow at a faster rate just after the professional development experience (during the 1st year of participation) and then slow down or even reverse (i.e., become negative) in subsequent years. Exploratory analyses confirmed that this was the case and, consequently, the level one model was specified to permit the estimation of two separate growth rates. This is commonly referred to as a piecewise growth model (see Bryk & Raudenbush, 1992; Seltzer, Frank, & Bryk, 1994).

RESULTS

The overall picture that emerges from modeling science and mathematics teachers' growth in their attitudes, preparation, and inquiry-based practice is one of short-term growth and long-term stability. The attitudes, preparation, and practice of both science and mathematics teachers showed substantial and statistically significant gains from before their summer professional development to the following spring. These gains ranged from one half to a full standard deviation in magnitude. In the second time period, from the end of the 1st year to the end of the 3rd year, teachers exhibited either slight growth or decline, on average, but these were generally not statistically significant changes. The following sections discuss the findings of the science and mathematics models in greater detail.

Table 4 shows the coefficients for the parameters used to predict teachers' growth in science and mathematics attitudes, preparation, and practice. The first set of coefficients is used to predict the initial status of teachers' attitudes, preparation, and practice before the professional development activities. The second set of coefficients predicts teachers' growth in attitudes, preparation, and practice 1 year after the professional development. The third set of coefficients predicts teachers' growth in attitudes, preparation, and practice up to 3 years later.

Preprofessional Development

The intercepts at initial status are negative simply because, on average, teachers started significantly below average (i.e., average for the duration of the study) on the measures of attitudes, preparation, and practice relative to where they ended up 4 years later. This was not surprising given the 4-year time period measured in this study. Interestingly, there was no significant relationship between a school's climate and teachers' initial attitudes toward inquiry. But teachers at schools with stronger reform climates tended to have higher initial preparation for inquiry and started their professional development using inquiry-based teaching practices with greater frequency. Surprisingly, teachers at schools with higher proportions of minority students tended to have higher levels of initial preparation for inquiry-based science. There were no differences in the initial statuses of public school teachers compared with private school teachers.

Several individual teacher characteristics were significant predictors of teachers' initial attitudes, preparation, and practice in both science and mathematics. First, teachers' prior professional involvement was a strong predictor of their initial attitudes, preparation, and practice; in other words, teachers with more professional involvement tended to have higher levels of attitudes, preparation, and inquiry-based instructional practices than those with less involvement. Second, teaching experience was not associated with teachers' attitudes toward inquiry-based science or mathematics or their use of inquiry-based teaching practices. In fact, teachers with more years of experience tended to feel less prepared to use both inquiry-based science and mathematics instructional practices initially than did teachers with less experience; every additional year of teaching experience was associated with approximately one tenth of a standard deviation smaller initial preparation level. This

Table 4

Coefficients for Individual and School Characteristics in Two-Piece Growth Models of Teachers' Attitudes and Preparation to Use Science and Mathematics Reform and Use of Mathematics Reform Teaching Practices

	Science			*Mathematics*		
	Attitudes	*Preparation*	*Practice*	*Attitudes*	*Preparation*	*Practice*
Initial status						
Intercept	–.444****	–.681****	–.596****	–.630***	–.390**	–.527**
School climate	.010	.347****	.341****	.044	.315****	.315****
Percentage of minority students in school	–.073	.429***	–.047	–.059	.451***	–.268*
Public school teacher	–.113	.213*	.044	.168	–.076	.116
Teacher involvement in reform	.130****	.103****	.108****	.182****	.122****	.190****
Years of teaching experience	–.002	–.011***	–.001	.003	–.014****	.006
Female teacher	.377****	.019	.236***	.274***	–.020	.081
Elementary school teacher	.218***	.051	.220***	.014	.095	.013
High school teacher	–.210***	–.230***	–.344****	–.256**	–.166*	–.128
African American teacher	–.168	.406**	.384***	–.347**	.318**	.146
Growth to 1 year						
Intercept	.504***	.731****	1.017****	.567****	.658****	.551***
School climate	.017	–.202***	–.225****	–.054	–.112	–.240***
Percentage of minority students in school	–.037	–.348**	.021	.233	–.252	.063
Public school teacher	.035	.159	–.233*	–.018	.116	.151
Years of teaching experience	.001	.008*	.001	–.001	.009*	–.003
Female teacher	.063	.132	.078	.190*	.218**	.109
Elementary school teacher	.028	.114	–.005	.162	.002	–.079
High school teacher	–.094	.071	.055	–.125	.134	–.034
African American teacher	.057	–.646***	–.390**	–.203	–.163	–.242

Growth from 1 to 3 years						
Intercept	–.129	.120	–.324**	–.158	–.020	.270
School climate	–.122	.028	–.042	.052	–.025	–.008
Percentage of minority students in school	.047	.060	–.087	.085	.244	.331**
Public school teacher	.243	–.241	.198	.116	.125	–.115
Years of teaching experience	–.001	–.006	–.002	–.005	–.001	–.002
Female teacher	–.061	.090	.123	–.048	–.059	–.159
Elementary school teacher	.013	.033	–.125	.023	–.003	.094
High school teacher	.215*	.150	–.026	.019	–.033	.001
African American teacher	–.031	.610**	.060	.495*	–.233	–.091

*$p \leq .10$. **$p \leq .05$. ***$p \leq .01$. ****$p \leq .001$.

suggests that more experienced teachers tend to feel less comfortable with more reform-oriented teaching practices.

Third, female teachers had significantly more positive attitudes toward inquiry-based science and mathematics teaching than did male teachers. Female teachers used inquiry-based teaching practices more frequently in science (but not in mathematics) than their male counterparts. Fourth, elementary school teachers of science had more positive starting attitudes and higher levels of practice than did middle school teachers (the omitted and therefore referent group); both elementary and middle school teachers had higher initial values of attitudes, preparation, and practice than did high school teachers. There were no differences between elementary and middle school teachers of mathematics.

Finally, African American teachers of science, who represented 7% of the sample of teachers, had higher levels of preparation for inquiry-based science teaching and used inquiry-based science instructional practices with greater frequency than did White teachers. African American teachers of mathematics had approximately one third of a standard deviation lower attitudes than did White teachers, but approximately one third of a standard deviation stronger initial preparation for reform than did White teachers.

Growth to 1 Year

Even after controlling for the differences in initial status, teachers' attitudes, preparation, and practices strongly increased after their involvement in both Discovery's science and mathematics professional development. At the end of the first growth period, teachers' attitudes, preparation, and use of inquiry-based instructional practices in science and mathematics grew by one half to a full standard deviation, on average.

By and large, the characteristics of the teachers and their schools were not related to their growth rates. Growth did not differ significantly for teachers in public versus private schools, for teachers with different levels of teaching experience, or for elementary, middle, or high school teachers. Teachers in schools with high minority populations also grew similarly to teachers in low minority population schools (except for feelings of science preparation). Although they started at a higher point in terms of practice, teachers who worked in schools with a more reform-oriented climate tended to have slower growth in teaching practice in the first time period than teachers who worked in schools with more traditional climates. This could be due to their higher initial scores on these dimensions.

Finally, in the first growth period, the preparation of African American teachers of science grew at a slower rate compared with their White counterparts. In fact, after the 1st year, the preparation of African American teachers had grown hardly at all, just eight hundredths of one standard deviation. Likewise, African American teachers reported that they changed their teaching practice at a slower rate than did White teachers, with African American teachers' teaching practices changing four tenths of a standard deviation less than their White counterparts. In mathematics, there were no differences between the growth patterns of different ethnic groups.

Growth From 1 to 3 Years

In the second time period, from a full year after the professional development to up to 3 years afterwards, gains in teachers' attitudes, preparation, and teaching practice were sustained. In other words, teachers' growth patterns flattened out, neither increasing nor decreasing. This can be seen by viewing the coefficients at the intercept of the second time period; there are nonsignificant dips in teachers' attitudes toward inquiry and preparation for inquiry, and there is a small but not significant gain in inquiry-based teaching practice. The exception was teachers who received professional development in science who reported significant (approximately one third of a standard deviation) declines in their inquiry-based teaching practices. This decline must be tempered by the recognition that science teaching practices showed the largest gain after the first time period.

Interestingly, after starting at an initially lower level of use of inquiry-based teaching practices, teachers in high minority schools tended to grow faster in their use of inquiry-based teaching practices than teachers in schools with fewer minorities. This suggests that changes in inquiry-based teaching practice take longer to occur in schools with higher proportions of minority students but follow the same general growth pattern. Finally, the slow growth of African American teachers' preparation for inquiry-based science teaching after the first time period can be viewed as a lag effect because African American teachers' preparation for inquiry grew significantly faster than that of White teachers in the second time period.

Figure 1 shows three panels for each subject that graphically depict the average teacher's change over time in his or her attitudes toward inquiry, preparation to use inquiry, and use of inquiry-based instructional practices. The graphs show the predicted values for White female middle school teachers with average involvement in reform and average teaching experience in schools with the average percentage of minority students and average school climates.

Two things stand out in these graphs. First is the positive and significant slope in the growth curves from before participating in Discovery until 1 year after participating. This shows the large positive gains that teachers made in all three areas during the time of the professional development experience.

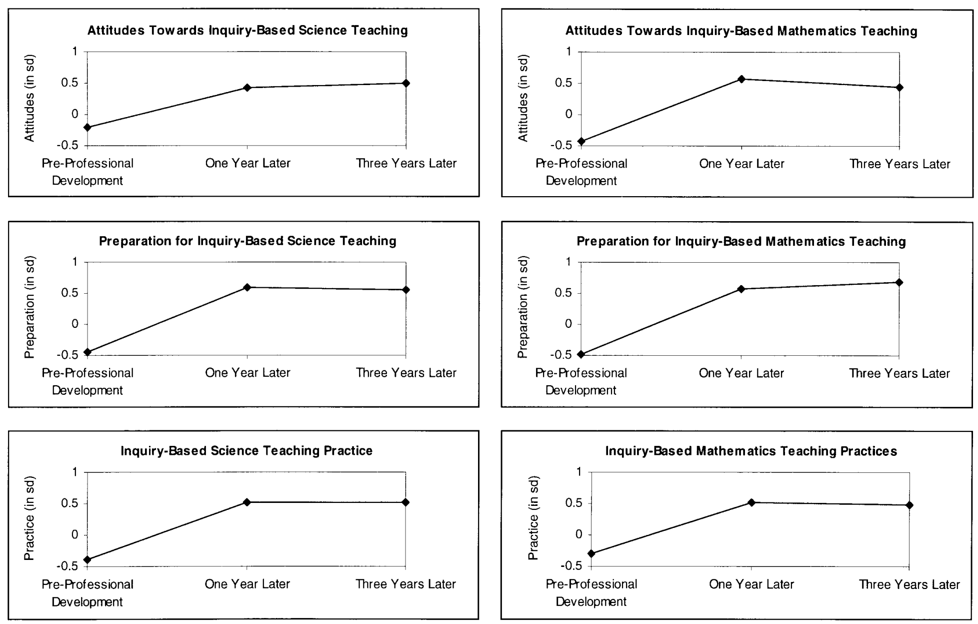

Figure 1. Change in Teachers' Attitudes, Preparation, and Use of Reform Science and Mathematics Teaching Practices
Note. Graphs represent predicted change over time for White female middle school teachers with an average involvement in reform and average teaching experience in schools with average school climate and percentage of minority students.

Second, these gains are sustained; from 1 to 3 years after their training, teachers continue to exhibit the same high levels of attitudes, preparation, and, most important, practice. Only the measure of teachers' attitudes declines, but the net effect is still a gain of two thirds of a standard deviation.

IMPLICATIONS

Together, the analyses of Ohio science and mathematics teachers provide evidence that suggests that Discovery's professional development in the context of Ohio's statewide systemic reform initiative may be associated with the widespread use of the inquiry-based instructional approach in the middle grades in the state. Teachers in both disciplines who participated in Discovery's intensive professional development showed substantial and statistically significant growth from before their training to a year later in their attitudes toward inquiry, their preparation to use inquiry-based pedagogy, and their actual use of inquiry-based teaching practices. Equally important, teachers' growth in these three areas was sustained during the 3 years following their professional development experience. Furthermore, Ohio's professional development appeared to have a similar impact on all teachers, regardless of their individual or school characteristics. The growth patterns generally were uniform across teachers of different genders, ethnic groups, grade levels, and school types (public vs. private).

The two variables that can be viewed as indicators of systemic reform—school climate and teachers' involvement in other professional teaching activities—yielded important clues about the environment within which reform was undertaken. Notably, school climate was influential in teachers' initial feelings of preparation to conduct inquiry and their actual use of inquiry-based teaching practices but not on their attitudes toward inquiry. However, the impact of school climate on preparation and teaching practice seemed to even out over time. This indicates two things. First, that teachers held their attitudes regardless of school context. Second, that school environment factors had a powerful influence on teachers' initial feelings of empowerment but that climate became less of an influence over time. We hypothesize from this that both the sustained nature of Discovery's professional development and the systemic support structures woven through it helped teachers overcome the particularities of their school's climate.

Another important finding related to the role of a systemic context is that teachers with prior involvement in professional teaching activities associated with systemic reform (attending association meetings, serving on curriculum or textbook committees, etc.) tended to exhibit higher initial attitudes, preparation, and practices. It is a weakness of our study that we measured professional involvement only at the initial status point and, therefore, were unable to model its influence over time. Further analyses should explore the different facets of systemic support in greater depth, allowing for finer-tuned distinctions and a better understanding of interactions between them.

In *Getting to Scale With Good Educational Practice* (1996), Harvard's Richard Elmore argued that "innovations that require large changes in the core of educational practice seldom penetrate more than a small fraction of U.S. schools and classrooms, and seldom last for very long when they do" (pp. 1-2). Elmore defined *core of educational practice* as the fundamental relationships between knowledge, teachers, and students. Inquiry-based instruction attempts to alter this core. From Elmore's perspective, both the Progressive reform movement of the 1920s and the curriculum reform movement of the 1950s were attempts at large-scale reform that failed.

Results from this study indicate that systemic reform sites of the 1990s that employ intensive professional development programs like those used in Ohio may have a different fate. Why may the third time be more charmed? What distinguishes the systemic reform movement from its predecessors in the 1920s and 1950s? Three factors may be contributing to differences in these efforts.

First, the systemic reform movement seems to have learned some important lessons from its antecedents about what comprises high-quality professional development: content-rich, intensive, sustained experiences that explicitly model the forms of teaching they intend participants to emulate. Additionally, the mantra of high standards has offset potentially weakening charges that the curriculum is being dumbed-down and simultaneous emphases on equity has helped keep the reforms away from the equity or excellence pendulum that seems to sweep the country every decade or so. Finally, the forefront emphasis on high-quality professional development, with standards-based curricula in the background, has helped to avoid the teacher-proof curricula problems that beset the large-scale curriculum development projects of the 1950s.

The second factor that may contribute to this movement being more effective is the foundation of support provided by the standards movement. Standards, implemented through both federal legislation (e.g., Goals 2000 and the reauthorization of the Elementary and Secondary Education Act), and state and national documents (e.g., professional, content, and performance standards) provide a shared vision of effective professional development of teachers, instructional methods, and conscious links to other parts of the education system (Loucks-Horsley, Stiles, & Hewson, 1996). Although several questions remain as to how standards at different levels interplay (see Darling-Hammond, 1994; Eisner, 1995) and how standards-based reforms are translated into real policies (see Massell, Kirst, & Hoppe, 1997), there can be little doubt that standards are having a powerful influence on the thinking of education reformers.

Third, the systemic reform model provides an important framework within which professional development takes place. Although there are several strains of systemic educational reform (see Vinovskis, 1996), the underlying concept of systemic reform is to develop coherent educational policies through coordinated governance structures that seek to align the various components of the educational system (i.e., standards, assessments, curricula, professional development, etc.) to provide reinforcing and complementing policies. Teacher learning is seen as a critical element of the implementation of systemic educational policy (Fullan, 1991) and a central component of systemic school restructuring (Elmore, Peterson, & McCarthey, 1996; Smylie, 1994). In Ohio, the SSI was supported by both the Board of Regents and the Department of Education, providing a coordinated governance system for the reform. In addition, enhancing the content knowledge of teachers and their skills in teaching by inquiry were critical elements in the Ohio reform. Furthermore, systemic reform researchers argue that a complement of top-down and bottom-up implementation methods limit the weaknesses of either a solely mandated or grassroots approach (Shields, Marsh, & Adelman, 1998). Ohio's regional infrastructure, coordinated from a central Discovery office, combined these two elements effectively. Regional councils had autonomy in selecting the academic leadership teams and in deciding how to allocate their funds. However, the type and length of professional development experiences, which evolved during the reform, were standardized across the state.

Despite these strengths, several questions remain to be addressed before it can be concluded that inquiry-based professional development, supported by standards-driven systemic reform, is the key element to changing the core of learning. One crucial question is whether inquiry-based instruction leads to gains in student learning. To answer this question, Discovery developed its own test and, using matched samples of Discovery and non-Discovery teachers, demonstrated a significant relationship between student performance on the test and teacher participation in the project's professional development. As the project progressed (Discovery continues to be funded by the Ohio General Assembly and is in its 9th year), Discovery has been able to assess changes in passing rates on Ohio's high-stakes proficiency tests in mathematics and science for students of Discovery and non-Discovery teachers. Passing rates on the proficiency tests in urban districts show more positive gains for students whose teachers have participated in Discovery compared with students of non-Discovery teachers. These results suggest that intensive professional development within a systemic framework can translate into student gains.

Another critical question surrounds the necessary intensity of training for different populations of teachers. Our analysis focused on an intensive, 6-week professional development model that was provided to about one fifth of the middle school teachers in Ohio who volunteered to participate. Is this concentration of training sustainable as it is scaled up throughout a system? Would similar impacts be seen with less intensive programs? Would the effects be the same for those teachers more reticent to volunteer? What are the optimal levels of professional development for teachers with different backgrounds, at different grade levels, and from different contexts?

In the 3rd year of its 5-year SSI grant, Discovery introduced several other types of professional development experiences. The 6-week institutes were trialed as two 3-week sessions (over 2 years) and as a 4-week institute. Furthermore, 40-hour-long workshops that were taught by Discovery teachers in their local districts were used to attract other teachers. In addition, because Discovery did not develop or provide curriculum materials, 2-week institutes, developed and taught by national curriculum groups, were introduced at the district level. Data concerning these alternative experiences have been collected as part of Discovery's ongoing assessment, but the results have not been analyzed.

Further explorations of the data from Ohio's SSI and other inquiry-based professional development programs would help to provide further evidence about the link between inquiry-based instructional practice and student achievement and help to tease out the effects of different intensities of training. But the findings of this study provide a promising indication that large-scale, high-quality, inquiry-based professional development set within a context of standards-based systemic reform can be a powerful mechanism for sustained and positive impact on teachers' attitudes, preparation, and teaching practices.

APPENDIX
Survey Questions and Scales Used to Create Mathematics and Science Composites

The following survey questions were used to create a composite of teachers' attitudes toward inquiry-based mathematics or science practice. All questions were on a 5-point scale ranging from *strongly disagree* to *strongly agree*; (m) or (s) signifies whether this item was used in a mathematics and/or science composite.

1. I enjoy teaching mathematics/science (m/s).

2. I organize my curriculum around the textbook (responses reversed) (m/s).
3. Teachers should know the answers to most questions students ask about mathematics/science (responses reversed) (m/s).
4. Students should never leave mathematics/science class feeling confused or stuck (responses reversed) (m/s).
5. An important issue is not whether students' answers to any mathematics/science question is correct but whether students can explain their answer (m/s).
6. Some people are good at mathematics/science and some just are not (responses reversed) (m/s).
7. Good mathematics/science teachers show students the correct way to answer questions they will be tested on (responses reversed) (m/s).
8. In learning mathematics, students must master topics and skills at each level before going on to higher level tasks (responses reversed) (m).
9. The more mathematics "drill" problems students work on in a class period, the more they will learn (responses reversed) (m).

The following survey questions were used to create a composite of teachers' use of inquiry-based mathematics or science instructional practices. All questions were on a 5-point scale ranging from *not well prepared* to *very well prepared*; (m) or (s) signifies whether this item was used in a mathematics and/or science composite.

How well prepared do you feel to do each of the following?

1. Manage a class of students who are using hands-on/manipulative materials (m/s).
2. Use cooperative learning groups (m/s).
3. Implement inquiry or discovery learning (m/s).
4. Present the applications of mathematics/science concepts (m/s).
5. Phrase questions to encourage more open-ended investigations (m/s).
6. Use computers as an integral part of mathematics/science instruction (m/s).
7. Use calculators as an integral part of instruction (m).
8. Teach groups that are heterogeneous in ability (m/s).
9. Encourage participation of females in mathematics/science (m/s).
10. Encourage the participation of underrepresented minorities in mathematics/science (m/s).
11. Inform students of career opportunities in mathematics/science (m/s).
12. Use performance-based assessment (m/s).
13. Use portfolios to assess student progress (m/s).
14. Involve parents in the mathematics/science education of their children (m/s).

The following survey questions were used to create a composite of teachers' use of inquiry-based mathematics or science instructional practices. All questions were on a 5-point scale ranging from *never, once or twice a semester, once or twice a month, once or twice a week*, to *almost daily*; (m) or (s) signifies whether this item was used in a mathematics and/or science composite.

1. Work in pairs/teams/small groups (m/s).
2. Participate in a dialogue with the teacher to develop an idea (m/s).
3. Make conjectures and explore possible methods to solve a problem (m/s).
4. Do hands-on/manipulative activities (m/s).
5. Write their reasoning about how to solve a problem (m/s).
6. Work in class on a project that takes a week or more (m/s).
7. Use a computer (m/s).
8. Learn by inquiry (m/s).
9. Engage in reflective thinking/writing about what they learned (m/s).
11. Participate in science-related debate, role-play, or simulation (s).
10. Use calculators to explore problems (m).
11. Use calculators to do computations (m).
12. Use calculators to develop an understanding of mathematical concepts (m).

NOTES

1. Despite two mailings at the end of the third and fourth time periods to increase response rates, these rates were still fairly low, which is not surprising given that they were administered 2 and 3 years after the teachers' summer professional development experience. However, we were concerned as to whether the lower response rates at Times 3 and 4 would invalidate our ability to generalize these data to the population of participants. To test for nonresponse bias, we conducted a series of *t* tests, comparing the demographic (gender, ethnicity, school level, years of experience, and professional involvement) and school characteristics (public/private, percentage of minority students, school climate) of respondents at Times 3 and 4 to their initial populations. For the science teachers, there were no differences between Times 1 and 3 on any dimension. At Time 4, there was a different ethnic composition compared to Time 1 (a smaller proportion of African American respondents) and the teachers reported coming from schools with slightly better climates. Other than these few differences, there was no other evidence of differences between science respondents and nonrespondents. For the mathematics teachers, there was only one significant difference. At both Times 3 and 4, the respondents were slightly older (41 to 37 years, on average) than those at Time 1. However, the mathematics respondents did not differ from their counterparts on other background and school characteristics. Therefore, we decided to pursue the analyses for all four time periods.

2. Of the 781 science teachers and 694 mathematics teachers who completed surveys, approximately 10% either did not give their identification number, which is required for matching responses longitudinally, or their identification numbers were not accurately scanned. These teachers were dropped from the analysis. All subsequent data are reported for the 701 science teachers and 603 mathematics teachers for which we have identification numbers. Comparisons of those teachers retained and those dropped revealed few differences.

3. The number of Hispanic and Asian teachers who participated in the professional development was so small that they were omitted from the sample for the purposes of this analysis.

REFERENCES

Arons, A. B. (1989). What science should we teach? In the Biological Science Curriculum Study (Ed.), *A BSCS thirtieth anniversary symposium: Curriculum development for the year 2000* (pp. 13-20). Colorado Springs, CO: BSCS.

Atwater, M. M. (1994). Research on cultural diversity in the classroom. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 558-576). New York: Macmillan.

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

Bryk, A. S., Raudenbush, S. W., & Congdon, R. T. (1996). *Hierarchical linear and nonlinear modeling with HLM/2L and HLM/3L programs.* Chicago: Scientific Software International.

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

Bybee, R. W. (1993). *Reforming science education: Social perspectives and personal reflections.* New York: Teachers College Press.

Corcoran, T. B., Shields, P. M., & Zucker, A. A. (1998). *The SSIs and professional development for teachers*. Menlo Park, CA: SRI International.

Cremin, L. (1961). *The transformation of the American school*. New York: Knopf.

Cuban, L. (1984). *How teachers taught: Constancy and change in American classrooms, 1890-1980.* New York: Longman.

Darling-Hammond, L. (1994). National standards and assessments: Will they improve education? *American Journal of Education, 102*, 478-510.

Dewey J. (1915). *Schools of to-morrow.* New York: E. P. Dutton.

Dow, P. (1991). *Schoolhouse politics: Lessons from the Sputnik era*. Cambridge, MA: Harvard University Press.

Eisner, E. W. (1995). Standards for American schools: Help or hindrance? *Phi Delta Kappan, 76*(10), 758-764.

Elmore, R. F. (1993). *The development and implementation of large-scale curriculum reforms.* Cambridge, MA: Harvard Graduate School of Education. Center for Policy Research in Education.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*(1), 1-25.

Elmore, R. F., Peterson, P. L., & McCarthey, S. J. (1996). *Restructuring the classroom: Teaching, learning, and school organization.* San Francisco: Jossey-Bass.

Fennema, E. (1990). Teachers' beliefs and gender differences in mathematics. In E. Fennema & G. C. Leder (Eds.), *Mathematics and gender* (pp. 169-187). New York: Teachers College Press.

Fullan, M. (1991). *The new meaning of educational change.* New York: Teachers College Press.

Grobman, A. (1969). *The changing classroom: The role of the biological sciences curriculum study.* New York: Doubleday.

Kahle, J. B. (1996). Opportunities and obstacles: Science education in the schools. In C. S. Davis, A. Ginorio, C. Hollenshead, B. Lazarus, & P. Rayman (Eds.), *The equity equation: Fostering the advancement of women in the sciences, mathematics, and engineering* (pp. 57-95). San Francisco: Jossey-Bass.

Loucks-Horsley, S., Stiles, K., & Hewson, P. (1996). *Principles of effective professional development for mathematics and science education: A synthesis of standards.* Madison, WI: National Institute for Science Education.

Marder, C. (1996, July). *Evaluation of NSF's Statewide Systemic Initiatives (SSI) program: Compilation of state data from the annual report to NSD: SSI Program, 1995.* Menlo Park, CA: SRI International.

Massell, D., Kirst, M., & Hoppe, P. (1997). *Persistence and change: Standards-based systemic reform in nine states*. Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

Mayer, D. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, *21*(2), 29-45.

McDermott, L. C. (1996). *Physics by inquiry* (2 Vols.). New York: John Wiley.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Research Council (1996). *National science education standards.* Washington, DC: National Academy Press.

Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, J. L., & Schneider, S. A. (1993). *Reform up close: An analysis of high school mathematics and science classrooms.* Madison: Wisconsin Center for Educational Research.

Rogosa, D., & Saner, H. (1995). Longitudinal data analysis examples with random coefficient models. *Journal of Educational and Behavioral Statistics*, *20*(2), 149-170.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, *90*, 726-748.

Seltzer, M. H., Frank, K. A., & Bryk, A. S. (1994). The metric matters: The sensitivity of conclusions about growth in student achievement to choice of metric. *Educational Evaluation and Policy Analysis*, *16*(1), 41-49.

Shields, P. M., Marsh, J., & Adelman, N. E. (1998). *The SSI's impacts on classroom practice*. Menlo Park, CA: SRI International.

Smylie, M. A. (1994). Redesigning teachers' work: Connections to the classroom. *Review of Research in Education*, *20*, 129-177.

Vinovskis, M. A. (1996). An analysis of the concept and uses of systemic educational reform. *American Educational Research Journal*, *33*(1), 53-85.

Weiss, I. R., Matti, M. C., & Smith, P. S. (1994). *Report of the 1993 National Survey of Science and Mathematics Education*. Chapel Hill, NC: Horizon Research.

Willett, J. B. (1994). Measuring change more effectively by modeling individual change over time. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed.). Oxford, UK: Pergamon.

Zucker, A. A., Shields, P. M., Adelman, N., & Powell, J. (1995). *Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) program: Second-year report. Part I: Cross-cutting themes* (SRI Project 3612). Menlo Park, CA: SRI International.