Is Alignment Enough? Investigating the Effects of State Policies and Professional Development on Science Curriculum Implementation

WILLIAM PENUEL
SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025, USA

BARRY J. FISHMAN
The University of Michigan, 610 East University, Room 1360B, Ann Arbor, MI 48109-1259, USA

LAWRENCE P. GALLAGHER, CHRISTINE KORBAK, BLADIMIR LOPEZ-PRADO
SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025, USA

ABSTRACT: Implementation of science curriculum materials has been a fundamental challenge in science education for decades. Policy researchers have argued that alignment of standards, curriculum, and assessment are the key to supporting implementation. This paper focuses on teachers' perceptions of curricular alignment and on curriculum implementation using empirical data from a statewide systemic inquiry science reform effort targeting students from kindergarten to eighth grade. We find that the success of alignment policies depends on teachers' construal of the relationship between standards and curriculum materials and on allocation of time for planning at the school level. © 2008 Wiley Periodicals, Inc. *Sci Ed* **93:**656–677, 2009

INTRODUCTION

Curriculum implementation in science is a problem that has long vexed policy makers, curriculum developers, and science educators. In the late 1950s, when the National Science Foundation first funded the design of hands-on science materials for schools, curriculum developers became frustrated by what they saw as teachers' failure to enact curricula in ways that reflected an understanding of the structure of scientific disciplines (Bruner, 1960). More recently, learning sciences researchers engaged in curriculum development projects have found that teachers using reform-oriented curriculum materials enact them either to a limited extent or in ways that do not reflect the intentions of designers (Brown & Campione, 1996; Reiser et al., 2000; Songer, Lee, & Kam, 2002; Spillane, 1999). These observations about science curriculum implementation are similar to observations made about reading and mathematics programs, where concern about implementation and the consequences of poor implementation on impacts on student outcomes are recurring themes in evaluation studies (Rowan, Camburn, & Correnti, 2004; Sarama, Clements, & Henry, 1998).

As a solution to the problem of implementation in other subject areas, policy makers have sought to promote greater alignment among standards, curriculum, assessment, and professional development (Fuhrman, 1993a, 1993b; M. Smith & O'Day, 1991). The assumption behind this strategy is that without good alignment among standards, curriculum, and assessments, teachers face a conflicting configuration of demands on their practice that is difficult for them to interpret and act upon (Fuhrman, 1993b). Including professional development as part of a broader strategy of aligning standards, curriculum, and assessments has emerged as a key policy instrument, since teachers need opportunities to learn about what new standards and curricula demand of them so that they interpret policy demands accurately and enact materials as intended (Hill, 2001).

This paper focuses on a single state's attempt to support implementation of an inquiry-oriented science curriculum by aligning standards, curriculum, assessment, and resources for professional development. The context was ideal for studying the potential effects of alignment, since the state's approach to promoting teacher learning about its policies was comprehensive and closely aligned to emerging best practice for teacher professional development, and the state had devoted significant levels of resources to support implementation. Furthermore, although the state in question did not have a science test integrated into its accountability system, the program itself had a system for monitoring implementation in the classroom, a precondition for tight coupling between policy and practice (H.-D. Meyer & Rowan, 2006). But even with all of these efforts at creating alignment, implementation rates of the curriculum in this state were disappointingly low from policy makers' perspectives.

This paper answers the following question: On what teacher characteristics and aspects of the local setting was the success of alignment as a policy instrument in promoting curriculum implementation contingent? Using survey methods, we analyzed the overall level of curriculum implementation among teachers who participated in state-sponsored professional development activities, and we analyzed the influence of the state's policy tools and guidance on implementation. We fit multilevel quantitative models of surveys that indicated that the policy tools the state used had little effect on implementation or on teachers' perceptions of alignment. Instead, implementation was related to the availability of multiple curriculum materials in teachers' schools and to differences between the state's and teachers' construal of what standards could be taught with the materials. Teachers' perceptions about how well aligned the materials were with their own and their schools' goals were related to these same factors and also to available time to plan and prepare for implementation and the level of accountability pressure on the school. These findings suggest potential limits to policy alignment as a strategy and to the need for greater consideration

of the larger ecology of influences on teachers' practices and to specific aspects of teacher knowledge. We believe that the issues raised by this study have implications for future curriculum development projects in science education.

BACKGROUND TO THE STUDY

One way to distinguish among policies in education is by the tools or instruments that policy makers use to effect change, and in recent years, states have sought to use tools aimed at systemic, rather than piecemeal, change (McDonnell & Elmore, 1987). Policies intended to effect systemic change require the coordinated use of multiple policy instruments aimed at multiple levels of complex school systems (M. Smith & O'Day, 1991). Among the most widespread of systemic change efforts in recent years have been those aimed at promoting greater *alignment* among standards, curriculum, assessments, and professional development. What alignment means to policy makers and to teachers, however, differs from state to state, from district to district, and from school to school. Furthermore, as with other kinds of policies, the success of alignment strategies in affecting teachers' perceptions and implementation of curriculum materials that policy makers promote is likely to depend on people and places where curricular enactment occurs (Honig, 2006). Below, we review the logic behind alignment as a policy instrument and evidence of success of the strategy in changing teacher cognition and practice, and we consider hypotheses about what could influence the success of alignment as a strategy for promoting science curriculum implementation that motivate the current study.

Alignment as a Policy Instrument for Promoting Curriculum Implementation

For many years, policy researchers pointed to a lack of alignment of curriculum, standards, and assessment as a cause of failure for particular policies intended to change teachers' practice. According to this argument, poor alignment among these three key elements of instructional policy can result in teachers having difficulty interpreting and acting upon the policy makers' demands (Fuhrman, 1993b). Furthermore, when teachers believe curricular activities will not prepare students to do well on assessments of student learning for which schools are held accountable, teachers may choose not to implement those materials (Li, Klahr, & Siler, 2006).

Improving alignment among standards, curriculum, and assessments at the state level is a system change strategy that policy makers have pursued in the last decade to improve curriculum implementation. Scholars argue that such alignment is necessary to produce clear policy guidance that teachers can follow (Herman & Webb, 2007; Knapp, 1997; M. Smith & O'Day, 1991). Accountability systems linked to assessments aligned with standards and curriculum materials provide sanctions and rewards intended to influence teachers' decisions about whether to follow policy guidance (Elmore, 2000; Herman & Webb, 2007). In addition, some federal funding streams in other subjects, such as Reading First, mandate that states select curriculum materials with particular features (e.g., focus on basic reading skills, have evidence of effectiveness) for teachers to use with students (Moss, Jacob, Boulay, Horst, & Poulos, 2006). These tighter controls have the potential to limit teachers' choices with respect to materials they use and to reduce heterogeneity in the curricular goals pursued by schools, which policy makers believe can increase chances for successful implementation.

A core component of effective alignment is the allocation and activation of resources to support teacher learning (Knapp & Plecki, 2001). Teachers need opportunities to learn

about the meanings of policies and their implications for practice (Cohen & Hill, 2000; McLaughlin, 2006). When states and districts fail to allocate sufficient resources to support teacher learning about policy aims and about the curricula designed to support those aims, teachers often misinterpret those aims (Cohen & Hill, 2001; Spillane & Jennings, 1997). Sustained, content-focused professional development is key to helping teachers gain practical knowledge of specific curricular activities and develop an understanding of what classroom instruction with the materials should look like (Ball & Cohen, 1999).

Although researchers in science education policy have advocated improved alignment for many years (see, especially, Knapp, 1997), most states have made significant progress only in the area of developing student standards. In science, most textbooks in use provide limited opportunities to develop deep understandings of the concepts that are typically part of standards documents (Kesidou & Roseman, 2002). Examples of coherent science curricula that are widely available and that reflect an understanding of how to develop deep understanding are just now emerging (Kali, Linn, & Roseman, 2008; Krajcik & Reiser, 2004). State assessment systems in science tend to test less frequently than in other subject areas (every 3 years, as opposed to every year), and the No Child Left Behind Act requires testing in science but does not require states to tie test results to rewards and sanctions for schools. As a consequence, some states do include science in determinations of whether schools are making Adequate Yearly Progress (AYP) toward achievement goals, but others do not.

So far, the evidence that recent efforts to align curricula, standards, and assessments can affect teachers' decision making and practice in science education is mixed. Some researchers have found strong effects of policy alignment on teacher knowledge and beliefs and on teachers' classroom practice (Levitt, 2001; Swanson & Stevenson, 2002), but others have found only modest effects (Blank, Porter, & Smithson, 2001). A recurring theme in research that examines effects of policies on cognition is that local actors often misinterpret the intentions of policy makers (Hill, 2001; Spillane, 2004; Spillane & Callahan, 2002). As local actors interact with one another, moreover, they develop shared but local meanings of policies and curricular interventions that can proliferate misunderstandings of policy intentions (Reiser et al., 2000; Spillane, Diamond, Walker, Halverson, & Jita, 2001).

How People and Places Could Affect the Success of Alignment Strategies

As with other policies in education, the success of alignment strategies is likely to be contingent on people responsible for enacting change and the unique configuration of demands and resources on those people in the places where they work (Datnow, Hubbard, & Mehan, 2002; Honig, 2006; Spillane & Burch, 2006; Spillane, Reiser, & Reimer, 2002). State education agencies exert considerable institutional pressure on school actors and have the power to set standards, select assessments, authorize the use of particular sets of curriculum materials, and allocate resources for professional development (Lusi, 1997). But teachers' prior knowledge and understanding of curricular purposes and standards can affect how they interpret the demands of policies on them and how they enact curricular materials, even when they receive professional development aimed at helping them change their practice. Furthermore, local actors face a unique configuration of state, district, and school pressures, and teachers may have access to curriculum materials that differ from those provided by the state as examples of aligned curriculum.

One source of teachers' misunderstandings of policy intentions is the knowledge they bring from their prior experience. For example, in a study of three reading teachers' responses to changes in their state's reading policies, Coburn (2004) found that teachers' responses differed depending on the kinds of instructional philosophies and strategies they

had encountered in their preservice education and in prior reform initiatives. The most common response of teachers to a change initiative in reading was for them to make modest changes that assimilated reform goals into frameworks and philosophies they had encountered earlier. Only when reform initiatives were coupled with intense and pervasive messages about their importance and meanings were teachers more likely to transform their practice, instead of making small, incremental changes to it.

With respect to alignment strategies, teachers' knowledge of standards and the relationship of standards to curriculum activities are likely to affect curriculum understanding. Past studies have shown that prior knowledge affects teachers' interpretation of state standards, which are often limited in the guidance they provide to teachers for how to interpret them (Hill, 2001, 2006; L. K. Smith & Southerland, 2007). Beyond understanding standards, teachers' understandings of the relationship between curriculum standards and purposes of curriculum may affect implementation. Comprehension of curricular purposes, as Shulman (1987) calls this form of knowledge, includes teachers' understanding of purposes that are embodied in policies, laws, standards, and other artifacts and given voice by educational leaders and fellow teachers (Magnusson, Krajcik, & Borko, 1999) and an understanding of how standards relate to big ideas in a discipline (Shulman, 1986). This form of knowledge, while difficult to measure, can be inferred from how teachers construe the relationship of standards to particular activities and in how that relationship is reflected in their decisions about what activities to implement to meet particular standards.

Teachers' decisions about what curriculum materials to use to address standards are based on more than just knowledge; they are also influenced by larger institutional dynamics. For example, most states leave decisions about what curriculum materials and instructional strategies to use to local districts and often also to teachers, even in the era of high-stakes accountability (Ingersoll, 2003). Furthermore, most teachers have access to a wide array of instructional materials made by different publishers in the private sector that reflect different and sometimes competing curricular purposes (H.-D. Meyer & Rowan, 2006). Some have suggested that these two institutional factors explain why some past systemic reform efforts met with limited success (Massell, Kirst, & Hoppe, 1997; Rowan, 2002), and it is likely that in alignment policies that still leave decisions about curriculum to teachers, those policies' effects may be diluted, depending on the availability to teachers of alternative curriculum materials.

Accountability systems themselves can also influence teachers' decisions about curriculum implementation, and for science, the likely effects are not well understood. The focus on reading and mathematics in federal and state policy has led school leaders to increase instructional time for those subjects, often at the expense of science instruction (Diamond & Spillane, 2004; Lee & Luykx, 2006; Marx et al., 2004). Even prior to passage of No Child Left Behind, case studies offer examples of schools where accountability pressure on elementary schools to increase achievement in reading and mathematics has diminished attention to science and science curriculum (Means et al., 2001; Spillane et al., 2001). In heightened accountability contexts, when states do devote attention to aligning curriculum, standards, and assessments in science, whether or not the increased attention mitigates accountability pressures in ways that can support curriculum implementation is not yet well understood. It may be that state-sponsored programs are more attractive to schools under pressure to improve test scores, because they carry the legitimacy of the state; but it may also be that schools' challenges in reading and mathematics continue to make science a difficult subject to which to devote precious instructional time.

Another setting that can affect the ways policies influence curriculum implementation is the school. School leaders are responsible for setting schedules for instruction and

for teachers' planning periods. Giving teachers time to meet during a regular part of the workday provides the conditions for coordinated efforts to implement curricular reforms (Elmore, 1996; Elmore, Peterson, & McCarthey, 1996). Specifically, it can facilitate teachers' engagement in making sense about the meaning of reforms (Coburn, 2001). Providing time for teachers to meet also increases the likelihood that they will have opportunities to learn from and share expertise with colleagues about instruction (Kruse, 2001). Shared planning time can positively influence teachers' perception of the school (Warren & Muth, 1995) and has been linked to classroom implementation of reforms (Supovitz & Turner, 2000; Uekawa, Aladjem, & Zhang, 2006).

Policy Tools Intended to Address Contingencies on People and Places

The dependency of policy implementation on particularities of people and places is often a worry for policy makers, but tools exist that policy makers have attempted to use to anticipate and address some of those particularities. Neither researchers nor policy makers need shy away from the complexity of the system they confront. The reality is that single-strategy approaches to policy making, even when they are as comprehensive as most efforts to improve alignment, rarely achieve their intended effects, often due to the complexities of the settings in which those strategies are implemented (Honig, 2006; Honig & Hatch, 2004). Alignment strategies may be strengthened by the use of policy instruments intended to address both individual teachers' situations and their organizational contexts.

One kind of policy instrument that is particularly relevant to addressing the reality of teacher agency is the use of strategies intended to persuade, rather than compel, teachers to implement particular curricula. Encouraging principal and district support for teachers' participation is an example of such a strategy that has the potential to persuade teachers to innovate and experiment (Scribner, Sawyer, Watson, & Myers, 2007). Another strategy for improving teachers' ownership of the implementation process is requiring them to vote as a faculty before adopting a particular program (Desimone, 2002). Although case studies in other disciplines point to the potential of these strategies for promoting curriculum implementation, we know little about what happens when states employ these strategies as part of a strategy of increasing alignment in science. Data from studies investigating states' strategies to secure principal support and teacher buy-in yield thus potentially important data on how effectively these strategies are in mitigating the effects of local control and individual teacher agency on curriculum implementation.

A strategy that policy makers have used to address teachers' need for better comprehension of curricular purposes is to provide teachers with maps showing the links among particular standards, curriculum activities, and the big ideas designers intended the curriculum to address. Both curriculum developers and states have developed such maps to both national standards (Quellmalz, Kreikemeier, Rosenquist, & Hinojosa, 2001) and to state standards (Kreikemeier & Quellmalz, 2002). The latter type of map is most likely to be salient to teacher decision making, not only providing teachers with information about what activities can be used to teach what standards, but also providing school actors at all levels with the necessary documentation to justify the use of the materials with students. Other scaffolds for teachers include materials that make clear the intentions behind particular activities and the relationship of activities to ideas that are central to a discipline.

With respect to accountability pressures, some states are exploring how to monitor implementation of science programs more closely and increasing accountability for results in science achievement. The intent of these strategies is to address the loose coupling that exists between science policy and teaching practice that derives from limited knowledge

among policy makers about what teachers are actually doing in science classrooms (J. W. Meyer & Rowan, 1977). In adding science testing to accountability systems, the expectation is that by incorporating science achievement into the system of rewards and punishments, schools will attend more closely to science than they have in recent years when the spotlight has been on performance in reading and mathematics.

Finally, policy makers can address school-level organizational contingencies by promoting the development of structures within schools that enable collaborative planning for instruction to occur among teachers. The strategies may support alignment policies by providing tools to help school actors reconcile policy goals with other, competing demands on schools for change (Honig, 2006). Such strategies have been promoted for several years by intermediary organizations in other subject domains, such as those that are part of the National Writing Project network (Lieberman & Wood, 2003). More recently, states and districts have been engaged in similar efforts to promote school-based collaborative learning groups focused on instructional matters. The Missouri Department of Elementary and Secondary Education, for instance, runs the Missouri Professional Learning Communities project, which provides resources and training to teams of teachers in schools who wish to engage in ongoing inquiry into teaching practice across a range of subject areas (Missouri Department of Elementary and Secondary Education, 2008).

Little empirical research exists on whether these tools, coupled with strategies to align standards, curriculum, assessments, and professional development, together can result in high levels of curriculum implementation. The popularity of the several tools described above—particularly alignment maps—suggests that such research is needed to inform policy. Furthermore, the continued movement toward greater accountability for science results makes this an opportune time to study the effects of these policy instruments on teacher cognition and practice, since accountability systems' success hinges on having a motivating effect on school actors. We undertook the current study, described below, to address these emerging policy and research needs.

THE CURRENT STUDY

The current study focuses on Alabama's implementation and perceptions of the GLOBE Program, an international program in Earth science and Earth science education. GLOBE provides curricular materials for use in classrooms, as well as an online database that is directed at supporting the work of scientists investigating aspects of the global environment. At the beginning of the program, scientists funded as GLOBE Principal Investigators by the National Science Foundation developed a set of protocols for students to use in four distinct investigation areas related to Earth systems: Atmosphere, Hydrology, Soils, and Land Cover/Biology. GLOBE "schools," which include a small number of other kinds of organizations such as science museums and senior centers, collect the data according to the protocols.

The implementation requirements for GLOBE schools focus on three critical elements. First, the program expects GLOBE teachers to implement the protocols for the program, using specialized equipment that can be purchased directly by teachers or sometimes is made available by GLOBE's professional development providers. In addition to collecting data, GLOBE expects teachers to report data to the GLOBE Web site for use by scientists in their own investigations (Means, 1998). Third, the GLOBE Program encourages teachers in classrooms to invite students to pose questions using student-collected GLOBE data as part of extended, local investigations of Earth systems (The GLOBE Program, 2005).

One of the places where GLOBE has been implemented with a great deal of attention is in the state of Alabama (Penuel, Shear, Korbak, & Sparrow, 2005), where the program

is implemented as part of the Alabama Mathematics, Science, and Technology Initiative (AMSTI). AMSTI is a statewide program funded by the Alabama Department of Education that was designed by a blue-ribbon panel to improve mathematics and science instruction across the state. The focus of AMSTI is on science in kindergarten through eighth grade. Besides GLOBE materials, other curriculum materials that are included as part of AMSTI training are science kits from the Full Options Science System (FOSS) program and from the Science and Technology Concepts (STC) program. FOSS and STC curriculum materials receive greater emphasis than do GLOBE materials in AMSTI, but officials stress that the GLOBE materials address key standards not addressed through the FOSS and STC kits (Nelson, personal communication, May 16, 2007). In addition, GLOBE protocols provide more direct encounters with the work of scientists than do the inquiry kits, in that scientists are expected to use GLOBE data from students in their own investigations.

The AMSTI initiative is an ideal context for studying the effects of alignment as a strategy for promoting curriculum implementation. First, AMSTI attempted to build district-level and school-level buy-in to its activities by selecting schools, rather than individual teachers, to be a part of AMSTI through a competitive application process, a strategy employed in many comprehensive school reform models (see, e.g., Cooper, Slavin, & Madden, 1998). As part of the application, schools must demonstrate that 80% of the mathematics and science faculty want to take part; if selected, all teachers from the school become participants in the initiative at no cost to the school. A second way that AMSTI promoted implementation was to make clear how GLOBE activities were aligned with state standards by providing teachers with clear instructional guidance (Cohen & Ball, 1999; Cohen & Hill, 2001; Rowan & Miller, 2007). The initiative provided teachers not only with specific GLOBE activities to implement in their classroom, but also with guidance on how and when to incorporate them into their instruction. Standards linkage documents show how specific GLOBE activities meet specific standards in each grade level. Third, AMSTI built in evaluation and feedback mechanisms for accountability purposes and for ongoing program improvement, which are essential features of contemporary standards-based reform models (Elmore, 2000). The program has published results from external evaluations on its Web site (Institute for Communication and Information Research, 2005) that document how student achievement in mathematics and science among AMSTI schools compares with the achievement in non-AMSTI schools. The program did not provide resources to promote the development of school-based groups to discuss implementation issues, although one idea behind having schools participate as entire faculty was to provide the conditions for such groups to form.

Another strength of the AMSTI design was its professional development model. AMSTI's activities to prepare teachers to implement GLOBE and other materials had characteristics associated with effective professional development (Desimone et al., 2002; Garet et al., 2001; Penuel, Shear, Korbak, & Sparrow, 2007). AMSTI professional development included a 2-week summer institute (during which teachers participate in 2 days related to GLOBE), equipment and materials, and follow-up support through specialists who are associated with regional science centers across the state. Third, AMSTI allocated social resources to support implementation in the form of a network of mentors, whom teachers could call on to answer questions and address problems of implementation.

Despite the strength of the policy and professional development supports for GLOBE implementation established through its integration into AMSTI, the data we present in this study show that the vast majority of teachers who took part in the initiative in the past year did not implement GLOBE protocols with students in their classrooms. Furthermore, although most teachers perceived the program as consistent with their own goals for professional development and with district and state expectations for student learning in science, nearly one fifth (17.7%) of teachers who responded to our questionnaire rated GLOBE materials

as "not sufficiently" consistent with these goals, even though the state took pains to identify the relationship between GLOBE and state standards. Both these results trouble AMSTI officials at the state level, both because they limit the potential impact of the initiative on achievement and because they call into question the efficacy of their efforts to create strong alignment of GLOBE to state standards and curriculum frameworks. As our models below suggest, however, much of what appears to be driving teachers' implementation are their perceptions of coherence at the individual level of materials alignment and their perceptions of local school context.

RESEARCH METHODS

Sample

The sample for the study was composed of teachers who attended AMSTI professional development institutes in summer 2006. All teachers who participated in one of three regional institutes in the state completed a background questionnaire in which we collected contact information from participants for later e-mail and mail follow-up. A total of 225 teachers from 51 different schools completed our full questionnaire.

Teachers in the study came from schools that were fairly representative of Alabama schools. A total of 29 schools in the sample or 57% had met criteria for AYP according to federal No Child Left Behind guidelines. That compares with 53% of schools statewide that met AYP. According to 2005 figures from the National Center for Education Statistics database, the student body for schools in our study was composed of 34.4% minority students (primarily African American) and 55.9% of students who were from low-income families, as measured by eligibility for the federal free and reduced-price lunch program. Statewide, 40.8% of students were from minority backgrounds and 43.0% of students were from low-income families. These figures for the final sample were similar to those for schools that completed only the initial survey: 48% of those schools had met AYP the prior year, 40% of students in the schools were minorities, and 49% of students served were from low-income families.

Roughly half of the teachers in the sample had a master's degree or higher, but only about one-fifth had teaching certificates in science (Table 1). The latter is partly explained by the sample composition, which consisted primarily of teachers who taught science at the elementary level. On average, teachers had 11.5 years of experience in the classroom ($SD = 8.8$ years) and 9.2 years of experience in teaching science ($SD = 7.7$ years). T-tests revealed that the sample of teachers and schools that completed surveys at baseline was similar to the final sample with respect to teachers' educational background, certification, and levels and years taught.

Sources of Data

We describe each of our sources of data below, describing data sources at both the teacher and school levels. For this study, the primary sources of data were a teacher questionnaire and the Alabama State Department of Education Web site. For all scales produced for analysis, we present reliabilities of those scales derived from our data set.

Dependent Variables

TEACHER Questionnaire: Alignment to Teacher and Organizational Goals.   We incorporated into our questionnaire a six-item factor ($\alpha = .93$) used in two earlier studies of

TABLE 1
Characteristics of Teachers

Characteristic	*n*	Percent
Highest level of education		
Bachelor's	94	42.3
Master's	110	49.5
Specialist	17	7.7
Doctorate	1	0.5
Science teaching certificates		
General science	19	8.5
Biology	14	6.3
Chemistry	1	0.5
Earth and space science	3	1.3
Physics	0	0.0
None	177	79.0
Teaching assignment for science		
Kindergarten	28	12.4
First	26	11.6
Second	32	14.2
Third	33	14.7
Fourth	30	13.3
Fifth	25	11.1
Sixth	20	8.9
Seventh	10	4.4
Eighth	12	5.3

professional development (Garet et al., 2001; Penuel et al., 2007), which measures how well the professional development matched the teacher's goals for professional development, the existing reform ideas within the school, and whether the professional development was followed up with activities that built upon what was already learned. In this study, we used alignment to teacher and organizational goals as a measure of the effects of different policy tools on teachers' perceptions of the importance of the GLOBE program.

Teacher Questionnaire: GLOBE Protocol Implementation. Use of GLOBE protocols with students was used as a measure of program implementation. Teachers indicated for each protocol whether they had implemented, planned to implement, or had not implemented the protocol and did not plan to do so. Using their responses to the questionnaire, we created an index that was a weighted sum of their intentions and actions across GLOBE's different investigation areas (1 unit for each protocol they intended to implement, 2 units for each protocol they implemented). The maximum value for the protocol implementation index was 14.

Teacher-Level Independent Variables

Perceptions of Alignment of GLOBE With Alabama Course of Study Science Content Topics. We focused our questions about teachers' perceived alignment of GLOBE with state content standards. We tailored items for each K–8 cluster of grade levels for which the standards are written—K–2, 3–5, and 6–8—and asked teachers to rate whether they covered the standard, using GLOBE or some other set of materials. We created an index of the sum of the number of standards teachers said they could meet using GLOBE materials

Science Education

for purposes of analysis. In addition to serving as a predictor in our quantitative models, comparing teachers' perceptions of alignment with those of the program developers (see below) provided us with a means to examine the effects of alignment strategies on teachers' decision making regarding curriculum.

Use of Alignment Tools and Artifacts From Partner. The GLOBE in Alabama partnership has developed an alignment tool that is posted on its Web site. (For first grade, see http://www.amsti.org/globe/documents/2008GLOBE/1st%20Grade%20Year%201% 20GLOBE/1st%20Grade%20Year%201%20GLOBE.pdf.) We asked teachers about their familiarity with this tool and also how/whether they had used it in making decisions about whether or how to implement GLOBE. Local AMSTI staff hypothesized that the use of these documents would aid teachers in making decisions about how to use GLOBE to meet state standards.

Barriers to Implementation. We asked teachers to identify the importance of each of seven types of barriers derived from earlier evaluation studies of the GLOBE program (Penuel et al., 2007). These barriers were difficulty finding time to prepare for implementing GLOBE, lack of technology access, lack of technical support for using computers and software, lack of GLOBE equipment, unsupportive school building administrators, and unsupportive district administrators. Teachers indicated whether each potential barrier to implementation was a major barrier, minor barrier, or not a barrier. For purposes of analysis, we analyzed barriers in three separate clusters. The first item, difficulty finding time to plan for implementation, was analyzed as a single categorical variable, recoding "major barriers" as 1 and "minor barrier" and "no barrier" as 0. The next three barriers related to equipment and technology had adequate reliability as a scale ($\alpha = .73$), so we analyzed them together. The last two items related to school and district support had good reliability for constructing a scale ($\alpha = .89$), so we analyzed those items as a single scale.

Support for Equipment Use. Because we found that partner support for equipment use was a significant predictor in our earlier professional development study of GLOBE implementation (Penuel et al., 2007), we included items used in our original study as dummy variables in the analysis. Equipment use support is a single survey item, which asked teachers whether the GLOBE partner provided assistance on technical setup and equipment use. This is a binary variable where $0 =$ no and $1 =$ yes.

Teacher Characteristics. We included items asking teachers about their certifications, education, and years of experience teaching, gender, and ethnicity as control variables in the analysis.

School-Level Independent Variable

State Assessment Data: Achievement Levels of Schools. We categorized schools in our sample on the basis of whether the school met AYP in both reading and mathematics for all grades and subgroups. AYP status was measured as a dichotomous variable in our analysis, and we intended to use it as an indicator of the effect of policy alignment on teachers' perceptions and GLOBE implementation. We anticipated that AYP status could affect implementation levels either by diverting attention away from science (since AYP status is based on mathematics and reading scores, not science scores) in the case of schools not meeting AYP, or that teachers in schools meeting AYP status might

report lower coherence and implementation levels, since they felt less pressure to adopt state-approved science curriculum materials.

Number of GLOBE Teachers in the School. We included the total number of GLOBE teachers in the school as a school-level variable. Since the policy was that teachers were expected to participate as part of a school team, we reasoned that the number of GLOBE teachers in the school might be positively related to perceptions of coherence and GLOBE implementation. In that respect, the variable can be interpreted as a social resource that schools allocated to support implementation.

Procedure

We pilot tested our questionnaire during the summer of 2006 using the technique of cognitive interviewing (Desimone & Le Floch, 2004). This technique involved having respondents "think aloud" as they answer questions on a paper questionnaire to ensure that respondents' thinking is aligned with the intent of particular items. On the basis of results from cognitive interviews, we revised items to reduce the likelihood that teachers would have trouble completing the questionnaire.

We then generated paper copies of the survey that could be scanned later for analysis. Using the initial questionnaire data collected from teachers at the time of training, we next sent the surveys to the 423 teachers on the rosters from the original survey. We then mailed questionnaires to each of those teachers in February 2007. We included an incentive fee of a $25 gift card for each teacher to complete the survey. Such fees, especially when given at the time questionnaires are mailed, can be effective in increasing response rates and reducing overall costs of follow-up.[1]

Two weeks after we mailed the questionnaires, we sent a postcard reminder to teachers to complete the survey. After 3 weeks, we sent reminders to all nonrespondents that included a cover letter from the AMSTI leaders at the Alabama Department of Education encouraging teachers to complete the survey. We conducted up to a total of 7 follow-up reminders (including e-mail and postal mail) to nonrespondents. The resulting response rate was 54%.

Once the data were collected, we scanned all forms and reviewed any discrepancies in data and response patterns. Standard quality control of the data was conducted for all indices and scales, such as checking for multicollinearity among variables in our models, skewness, and variance.

Data Analysis

We analyzed our data at two levels: teacher and school. Because of the nested structure of the data set, we used hierarchical linear modeling (Raudenbush & Bryk, 2002) to examine the relationship between effects of individual teachers' perceptions of alignment and use of alignment tools, interpretations of barriers, and experiences of support (Level 1) and the effects of school AYP status and number of GLOBE teachers in the school (Level 2) on teachers' global perceptions of the coherence of GLOBE and implementation of GLOBE. For each of our two outcome variables, coherence and implementation, we first conducted an unconditional model to determine the variance structure at each level (see Table 2). We then constructed a model for each outcome variable that analyzed the hypothesized links

---

[1]In our case, it is difficult to assess how effective the incentive fee was. Measuring the efficacy of an incentive fee would have required us to randomly assign teachers to either an incentive or no-incentive condition and compare response rates of teachers. In our study, such a test was beyond the resources of the study to support.

Science Education

TABLE 2
Variance Components: Teacher and School Levels

Level	Implementation		Global Perceptions of Coherence	
	Variance	Percent of Variance	Variance	Percent of Variance
Teacher (Level 1)	4.371	86.1	0.2919	91.2
School (Level 2)	0.7069	13.9	0.0283	8.8
Total	5.078	100	0.3202	100

between knowledge transfer processes and reform implementation. These models appear in the technical appendix.

RESULTS

Below we present descriptive statistics for each of the variables in the analytic models, followed by a description of the model results.

Descriptive Statistics

Perceptions of Alignment to Individual and Organizational Goals and Protocol Implementation. On average, teachers judged the alignment of GLOBE with teachers' own and their district's goals for student learning to be "sufficient," but implementation levels were low. On a scale from 0 to 3, with 0 being "not at all aligned" and 3 being "very well aligned," teachers in the sample rated the coherence of GLOBE as $M = 1.9$ ($SD = 0.57$, $n = 221$). With respect to implementation, roughly one quarter of the teachers reported that they had implemented Atmosphere protocols, and 14% reported they had implemented global positioning system (GPS) protocols. For the other investigation areas, less than 10% of teachers reported implementing any of these (see Figure 1). On average, teachers' scores on the index we constructed to measure protocol implementation were 2.6 ($SD = 2.2$, $n = 221$) out of a possible 14.

Coverage of Standards Using GLOBE and Use of Linking Documents. On average, teachers reported that they used GLOBE to cover 5.0 ($SD = 6.2$, $n = 221$) different content areas in the Alabama Course of Study, the state's content standards. About a quarter (26% or 56 of 213) of teachers, however, reported they used the linking documents provided by AMSTI to help guide their thinking about alignment to standards. Only 8.5% used these documents to make decisions about how to use GLOBE to replace existing materials in their curriculum.

A closer examination of teachers' reports of how they used GLOBE by standard reveals that teachers' decisions about what topics to use GLOBE materials to cover standards did not map well onto the AMSTI program's guidance. In many instances, only a small percentage of teachers used GLOBE to teach particular standards the program believed GLOBE could help teach. In other cases, teachers reported GLOBE could help them meet standards the program did not identify as linked to that standard. In Grade 3, for example, just 39% of teachers ($n = 10$) used GLOBE to teach about cloud types associated with particular weather patterns, even though the AMSTI program identified that as a standard that could be met by using GLOBE materials. By contrast, a similar percentage (31%) of
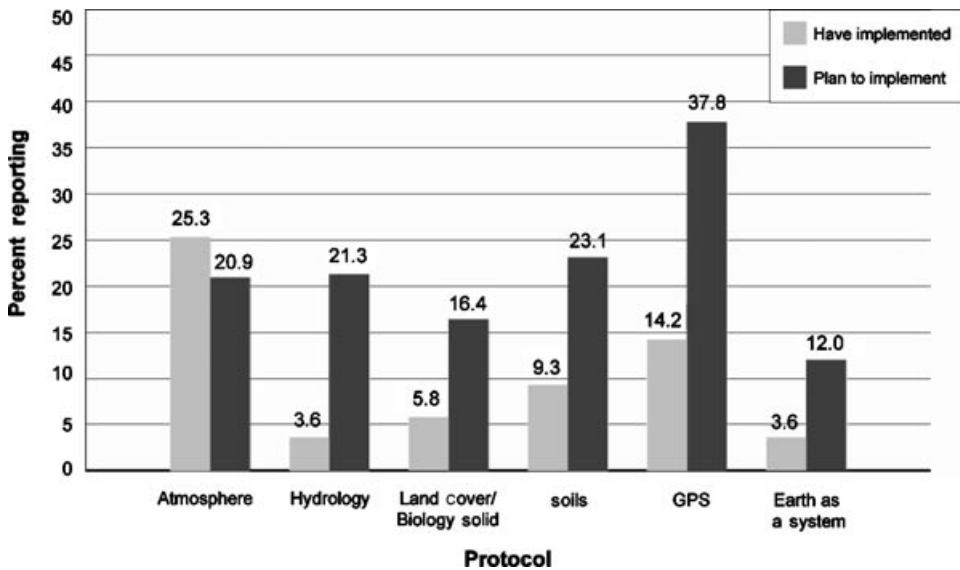
Figure 1. GLOBE implementation by investigation area.

third grade teachers said they used GLOBE materials to teach students to describe ways that energy from the sun is used, even though the program did not show a linkage between GLOBE and this standard (see Table 3). The patterns evident in Table 3 were evident across all grades; the table includes results for a single grade to illustrate the pattern.

TABLE 3
Teachers' Use of GLOBE to Cover Standards in Grade 3

| Standard	Emphasized in AMSTI PD for GLOBE	Cover Using GLOBE	Cover Using Other Materials
Describe ways energy from the sun is used		30.8% ($n=8$)	69.2% ($n=18$)
Define force and motion		4.0% ($n=1$)	96.0% ($n=24$)
Habitat conditions that support plant growth and survival		16.0% ($n=4$)	84.0% ($n=21$)
Identify cloud types associated with specific weather patterns	■	38.5% ($n=10$)	61.5% ($n=16$)
Identify positive and negative effects of weather phenomena	■	38.5% ($n=10$)	61.5% ($n=16$)
Identify technology used to record and predict whether	■	73.1% ($n=19$)	26.9% ($n=7$)
Explain symbols shown on a weather map	■	40.0% ($n=10$)	60.0% ($n=15$)
Organize weather data into tables or charts	■	52.2%	47.8%
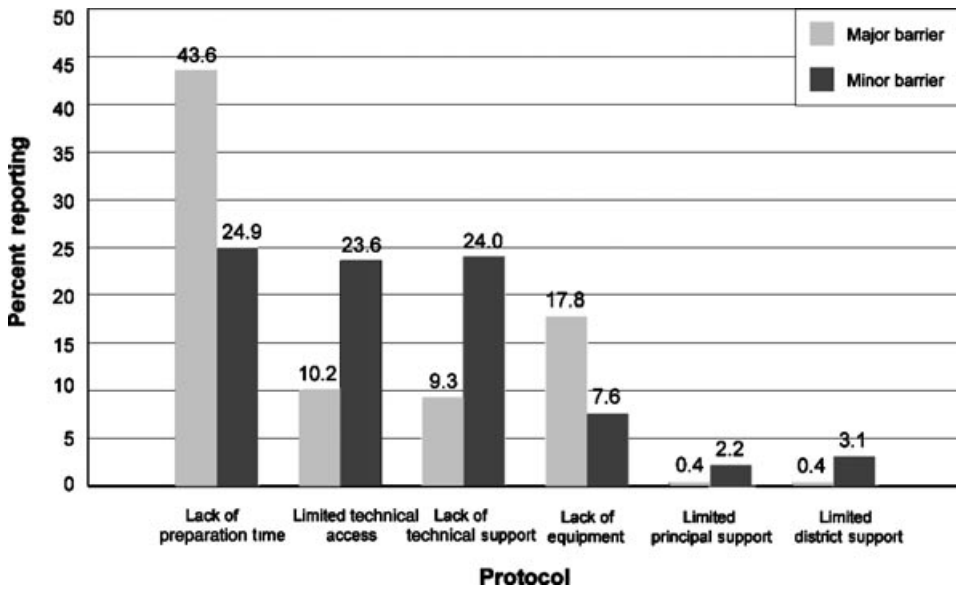
Figure 2. Barriers to implementation.

Barriers and Supports to Implementation

The most significant barriers to implementation reported by teachers were having inadequate time to prepare and plan for implementation and limited technology access and support (see Figure 2). Limited principal and district support were barriers to only a small percentage of teachers.

For the one support variable in our models, help with setting up GLOBE equipment, 36% ($n = 80$) reported they received this type of support.

School-Level Variables: AYP Status and Number of GLOBE Teachers in School. A total of 29 schools (57%) in the sample had met the criteria for AYP according to federal No Child Left Behind guidelines. There was a mean of 2.1 GLOBE teachers in the schools in the sample ($SD = 6.2$).

Predictors of Teachers' Perceptions of Alignment and Protocol Implementation

For the model testing the effects of the teacher- and school-level variables on teachers' perceptions of alignment to individual and organizational goals, the significant predictors were teachers' use of GLOBE materials to cover content standards, lack of time to prepare to implement GLOBE, and school AYP status (see Table 4). The more standards that teachers reported that they used GLOBE materials to cover, the higher their perceptions of alignment of the innovation with local goals for student learning. Conversely, when teachers lacked time to prepare, they reported less alignment. Teachers in schools that met AYP were less likely to view GLOBE as aligned with their local goals. None of the teacher background variables used as controls in the model was significant; therefore, we eliminated them from the models altogether.

Just one of the predictors included in our models predicted implementation levels of protocols (Table 4). Teachers who reported they used GLOBE to cover more standards

TABLE 4
Predictors of Teachers' Perceptions of Alignment and Protocol Implementation

Level/Variable	Perceived Alignment to Individual and Organizational Goals		Protocol Implementation	
	Coefficient	Standard Error	Coefficient	Standard Error
Teacher level				
Use of GLOBE to Cover Standards	.02[a]	0.01	.10[b]	0.03
Use of Standards Linking Documents	.07	0.07	.44	0.27
Lack of Time to Plan and Prepare	−.12[a]	0.05	−.05	0.23
Difficulties with Equipment and Technology Access	−.05	0.04	.09	0.17
Lack of principal and district support	.01	0.04	−.17	0.15
School level				
School AYP Status	−.18[a]	0.09	−.26	0.35
Number of GLOBE Teachers in School	.00	0.00	−.01	0.01
Intercept	2.07	0.12	2.25	0.50

[a]$p < .05.$ [b]$p < .001.$

were more likely to implement more protocols. This result is hardly surprising, since these two items are closely related. Teachers' use of linking documents was a significant predictor of implementation, but only at the $p = .12$ level.

DISCUSSION AND CONCLUSION

The results of this study suggest that there are at least some instances when alignment of curriculum materials with policies at the state level is not sufficient to promote a strong sense among teachers that particular curricular innovations cohere with their own goals or to promote widespread implementation of the innovation. The AMSTI initiative was a carefully crafted state-level science reform initiative, of which the GLOBE Program was an integral part. Efforts to persuade teachers of the alignment of GLOBE to their own and their school's goals by providing them with specific instructional guidance about how the program's curriculum materials aligned to state standards were largely unsuccessful, with teachers' judgments about the suitability of materials for teaching standards diverging widely from those of policy makers and professional developers in the program. Furthermore, despite efforts to promote a comprehensive approach to science reform in schools by inviting school-wide participation, the number of teachers participating from individual schools appeared to have little impact on teachers' perceptions of coherence or on their use of the GLOBE materials.

Similarly, the professional development had little impact on teachers' perceptions or on protocol implementation. Teachers' initial workshops included many opportunities for hands-on practice with implementing protocols; a system of mentors was also accessible for classroom-based support. Teachers also received all the equipment they would need to

implement their GLOBE protocols as part of the professional development, a factor we have found in past research (Penuel et al., 2007) can influence implementation of GLOBE. Still, teachers' ratings of the alignment of professional development to their own and their school's goals were modest, given these strengths, and their implementation levels were disappointingly (to program leaders) low.

Furthermore, none of the key strategies the state had employed to affect individual teacher decision making about the use of curriculum materials had a significant effect either on teachers' cognition or on their implementation. The alignment maps were not significant factors in teachers thinking or action. Thus, we found no evidence to support the hypothesis that efforts to address teacher knowledge or teacher agency in the state supported the overall alignment strategy of the state.

We argue that this case study indicates that one of the limits of strategies to promote greater alignment at the state or district level is that it fails to consider the particular needs of teachers and schools. Teachers' own reports of what GLOBE activities they used to teach standards reflected a misunderstanding of the standards that professional development would need to address. Furthermore, schools failed to allocate adequate time for teachers to prepare to implement the curriculum; despite their having clear instructional guidance, gathering and organizing materials for the program required time that teachers reported they were not given. In this respect, our findings are consistent with the perspective that alignment is an accomplishment that necessarily involves schools and cannot be achieved solely through state-level policy making (see, especially, Honig & Hatch, 2004).

The role of accountability pressure was significant as well, although its effect is not easily interpretable. When schools failed to meet AYP, teachers in those schools were more likely to perceive GLOBE as aligned with their goals. The reason could be that those schools were under pressure to follow state guidance such as is provided by AMSTI and that following such guidance improves local confidence in those schools. The explanation could also be that schools meeting AYP simply have more curricular resources and more discretion about how to use those than do schools not meeting AYP, leading fewer teachers to pick GLOBE materials as a way to teach particular standards.

Policy makers and program developers may conclude from these findings that teachers' choice in matters related to curriculum need to be reduced and that schools should be compelled to provide time for teachers to implement particular programs and curricula. But such an approach is inconsistent with a long-standing tradition in U.S. education of local control and with a desire to increase, rather than decrease, teacher professionalism by empowering them with the knowledge and resources they need to make good instructional decisions. As we see it, the challenge ahead for policy makers and program developers is to continue to work toward greater alignment at the state and district levels, while also developing strategies to help tailor professional development and implementation support in ways that advance multiple stakeholders' goals for science education.

Our findings suggest that one of the ways that professional development must be tailored is to develop more in-depth tools to support the development of teachers' comprehension of curricular purposes. Teachers may possess erroneous conceptions of the meaning of standards (Lin & Fishman, 2004), in which case the representations of alignment by curriculum developers are unlikely to be used in ways policy makers intend. Their judgments may also be affected by the sheer number of curricular resources from which they have to choose; their authority in being able to choose these materials means that professional developers promoting a *particular* curriculum face significant competition for teachers' attention and positive opinion. Variety in conjunction with poor understanding of the meaning of standards may combine to lead teachers to interpret thoughtfully designed and coherent curricula into sequences of activities that can be broken up and combined with

other district-provided materials, leading to incomplete or incoherent implementations of curriculum materials. There is some hope that recent curriculum development efforts, such as the Investigating and Questioning our World Through Science and Technology (IQWST) effort (Krajcik & Reiser, 2004), that aim to provide a comprehensive multiyear curriculum may avoid this problem by becoming the *sole* materials teachers are to use.

In addition to thinking more deeply about the kinds of resources that are needed to build teachers' comprehension of curricular purposes, state-sponsored models may also need to provide additional kinds of resources to schools to help guide them in thinking about how best to activate material and social resources to support implementation. Lack of planning time was a significant problem for many teachers in our study, and it was left up to schools to set schedules that allow teachers sufficient time to plan for curriculum implementation. But schools did not receive guidance about how much time might be needed to plan for implementation; in some ways, the alignment tools provided and materials implied it would be easy for teachers to simply take the materials they received and begin to implement immediately. In past studies, we have observed that GLOBE implementation is never so simple and has taken place only when teachers have the opportunity to share implementation challenges and successes with coaches and peers. Tools in the form of protocols for discussing implementation challenges or sharing effective activities might be useful to teachers for purposes of structuring conversations and expertise sharing.

An important limitation of this study is that it is a single test of our own interactive perspective on coherence in a single state context. Our sample, though representative of those teachers and schools participating in the GLOBE Program in Alabama, was not necessarily representative of teachers in the program more broadly, who are located in all 50 states. Different contexts, therefore, might yield different estimates of the predictors of coherence. What this study in one state does reveal, however, is that an exclusive focus on alignment at the state level or on teacher professional development models that do not take into account school-level constraints on teacher cognition and action can result in poor curriculum implementation. This finding is likely to be as sobering to policy makers, curriculum developers, professional developers, and researchers in contexts outside Alabama as it was to people closely affiliated with the GLOBE Program there. Furthermore, it is one more reminder of how the complexity of the educational system makes it difficult to effect broad changes in instructional practice.

TECHNICAL APPENDIX

We fit a two-level hierarchical linear model to the survey data. The Level 1 outcomes measured were teachers' perceptions of alignment to individual and organizational goals and GLOBE implementation. Below, we present the final model for teachers' perceptions of alignment (both models included the same predictors):

IMPLEMENTATION $_{ij} = \beta_{0j} + \beta_{1j}$ (STANDARDS COVERAGE$_{ij}$) $+ \beta_{2j}$ (USE OF LINKING DOCUMENTS$_{ij}$) $+ \beta_{3j}$ (LACK OF TIME$_{ij}$) $+ \beta_{4j}$ (EQUIPMENT BARRIERS$_{ij}$) $+ \beta_{5j}$ (LACK OF ADMIN SUPPORT$_{ij}$) $+ e_{ij}$

$$\gamma_{00j} = \gamma_{00} + \gamma_{01j} \text{ (AYP STATUS}_{1j}) + u_{0j}$$
$$\gamma_{10j} = \gamma_{10} + u_{1j}$$
$$\gamma_{20j} = \gamma_{20} + u_{2j}$$
$$\gamma_{30j} = \gamma_{30} + u_{3j}$$
$$\gamma_{40j} = \gamma_{40} + u_{4j}$$
$$\gamma_{50j} = \gamma_{50} + u_{5j}$$

where

$\beta_{0j} =$ intercept; the mean implementation level of teacher i in school j, controlling for all other variables included in the model;

$\beta_{1j} =$ the effect of the number of content standards GLOBE teachers used GLOBE activities to teach on protocol implementation;

$\beta_{2j} =$ the effect of teachers' use of state-provided linking documents showing connections between GLOBE activities and state standards on protocol implementation;

$\beta_{3j} =$ the effect of teachers' reporting that lack of time to plan was a barrier to implementation of GLOBE on protocol implementation;

$\beta_{4j} =$ the effect of teachers' reporting equipment access was a barrier on protocol implementation;

$\beta_{5j} =$ the effect of teachers' reporting that lack of administrative support was a barrier on protocol implementation;

$e_{ij} =$ error term associated with teacher i in school j;

$\gamma_{00} =$ school-level intercept; protocol implementation level for the average school, controlling for AYP status;

$\gamma_{01} =$ effect of a school's AYP status on protocol implementation; and $u_{0j} =$ error term associated with school j.

REFERENCES

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), Teaching as the learning profession (pp. 3–31). San Francisco: Jossey-Bass.

Blank, R., Porter, A. C., & Smithson, J. L. (2001). New tools for analyzing teaching, curriculum, and standards in mathematics and science: Results from the Survey of Enacted Curriculum Project Final Report. Washington, DC: Council of Chief State School Officers.

Brown, A. L., & Campione, J. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles and systems. In L. Schauble & R. Glaser (Eds.), Innovations in learning: New environments for education (pp. 289–325). Hillsdale, NJ: Erlbaum.

Bruner, J. S. (1960). The process of education. Cambridge, MA: Harvard University Press.

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. Educational Evaluation and Policy Analysis, 23(2), 145–170.

Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. Sociology of Education, 77(3), 211–244.

Cohen, D. K., & Ball, D. L. (1999). Instruction, capacity, and improvement. Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.

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.

Cooper, R., Slavin, R. E., & Madden, N. A. (1998). Success for all: Improving the quality of implementation of whole-school change through the use of a national reform network. Education and Urban Society, 30(2), 385–408.

Datnow, A., Hubbard, L., & Mehan, H. (2002). Extending educational reform: From one school to many. New York: Routledge/Falmer.

Desimone, L. M. (2002). How can comprehensive school reform models be successfully implemented? Review of Educational Research, 72(1), 433–479.

Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. Educational Evaluation and Policy Analysis, 26(1), 1–22.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. Educational Evaluation and Policy Analysis, 24(2), 81–112.

Diamond, J. B., & Spillane, J. P. (2004). High stakes accountability in urban elementary schools: Challenging or reproducing inequality? Teachers College Record, 106(6), 1145–1176.

Elmore, R. F. (1996). Getting to scale with good educational practice. Harvard Educational Review, 66, 1–26.

Elmore, R. F. (2000). Building a new structure for school leadership. New York: Albert Shanker Institute.

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

Fuhrman, S. H. (1993a). Designing coherent educational policy: Improving the system. San Francisco: Jossey-Bass.

Fuhrman, S. H. (1993b). The politics of coherence. In S. H. Fuhrman (Ed.), Designing coherent educational policy: Improving the system (pp. 1–34). San Francisco: Jossey-Bass.

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

Herman, J. L., & Webb, N. M. (2007). Alignment methodologies. Applied Measurement in Education, 20(1), 1–5.

Hill, H. C. (2001). Policy is not enough: Language and the interpretation of state standards. American Educational Research Journal, 38(2), 289–318.

Hill, H. C. (2006). Language matters: How characteristics of language complicate policy implementation. In M. I. Honig (Ed.), New directions in education policy implementation: Confronting complexity (pp. 65–82). Albany, NY: SUNY Press.

Honig, M. I. (2006). Complexity and policy implementation: Challenges and opportunities for the field. In M. I. Honig (Ed.), New directions in education policy implementation: Confronting complexity (pp. 1–23). Albany, NY: SUNY Press.

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. Educational Researcher, 33(8), 16–30.

Ingersoll, R. M. (2003). Who controls teachers' work? Power and accountability in America's schools. Cambridge, MA: Harvard University Press.

Institute for Communication and Information Research. (2005). Summary of AMSTI external evaluation, student achievement data, 2005. Tuscaloosa: Institute for Communication and Information Research, University of Alabama.

Kali, Y., Linn, M. C., & Roseman, J. E. (2008). Designing coherent science education: Implications for curriculum, instruction, and policy. New York: Teachers College Press.

Kesidou, S., & Roseman, J. E. (2002). How well do middle school science programs measure up? Findings from Project 2061's curriculum review. Journal of Research in Science Teaching, 39, 522–549.

Knapp, M. S. (1997). Between systemic reforms and the mathematics and science classroom: The dynamics of innovation, implementation, and professional learning. Review of Educational Research, 76(2), 227–266.

Knapp, M. S., & Plecki, M. L. (2001). Investing in the renewal of urban science teaching. Journal of Research in Science Teaching, 38(10), 1089–1100.

Krajcik, J., & Reiser, B. J. (2004). IQWST: Investigating and Questioning our World Through Science and Technology. Ann Arbor, MI: University of Michigan.

Kreikemeier, P., & Quellmalz, E. (2002). Aligning GLOBE to TEKS. Paper presented at the 4th Annual Texas Science Summit, San Antonio, TX.

Kruse, S. (2001). Creating communities of reform: Continuous improvement planning teams. Journal of Educational Administration, 39(4), 359–383.

Lee, O., & Luykx, A. (2006). Dilemmas in scaling up innovations in science instruction with nonmainstream elementary students. American Educational Research Journal, 42(5), 411–438.

Levitt, K. E. (2001). An analysis of elementary teachers' beliefs regarding the teaching and learning of science. Science Education, 86(1), 1–22.

Li, J., Klahr, D., & Siler, S. (2006). What lies beneath the science achievement gap: The challenges of aligning science instruction with standards and tests. Science Educator, 15(1), 1–12.

Lieberman, A., & Wood, D. (2003). Inside the National Writing Project: Connecting network learning and classroom teaching. New York: Teachers College Press.

Lin, H.-T., & Fishman, B. J. (2004). Supporting the scaling of innovations: Guiding teacher adaptation of materials by making implicit structures explicit. In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), Proceedings of the 6th International Conference of the Learning Sciences (p. 617). Santa Monica, CA: Erlbaum.

Lusi, S. F. (1997). The role of state departments of education in complex school reform. New York: Teachers College Press.

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge. In J. Gess-Newsome & N. G. Lederman (Eds.), Examining pedagogical content knowledge (pp. 95–132). Dordrecht, the Netherlands: Kluwer.

Marx, R. W., Blumenfeld, P. C., Krajcik, J., Fishman, B. J., Soloway, E., Geier, R., et al. (2004). Inquiry-based science in the middle grades: Assessment of learning in urban systemic reform. Journal of Research in Science Teaching, 41(10), 1063–1080.

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

McDonnell, L. M., & Elmore, R. F. (1987). Getting the job done: Alternative policy instruments. Educational Evaluation and Policy Analysis, 9(2), 133–152.

McLaughlin, M. W. (2006). Implementation research in education: Lessons learned, lingering questions and new opportunities. In M. I. Honig (Ed.), New directions in education policy implementation: Confronting complexity (pp. 209–228). Albany, NY: SUNY Press.

Means, B. (1998). Melding authentic science, technology, and inquiry-based teaching: Experiences of the GLOBE Program. Journal of Science Education and Technology, 7(1), 97–105.

Means, B., Penuel, W. R., Crawford, V. M., Korbak, C., Lewis, A., Murphy, R. F., et al. (2001). GLOBE Year 6 evaluation: Explaining variation in implementation. Menlo Park, CA: SRI International.

Meyer, H.-D., & Rowan, B. (2006). Institutional analysis and the study of education. In H.-D. Meyer & B. Rowan (Eds.), The new institutionalism in education (pp. 1–13). Albany: State University of New York Press.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340–363.

Missouri Department of Elementary and Secondary Education. (2008). Missouri professional learning communities. Jefferson City, MO: Author.

Moss, M., Jacob, R., Boulay, B., Horst, M., & Poulos, J. (2006). Reading First implementation evaluation: Interim report. Washington, DC: Office of Planning, Evaluation, and Policy Development, U.S. Department of Education.

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. American Educational Research Journal, 44(4), 921–958.

Penuel, W. R., Shear, L., Korbak, C., & Sparrow, E. (2005). The roles of regional partners in supporting an international Earth science education program. Science Education, 89(6), 956–979.

Quellmalz, E., Kreikemeier, P., Rosenquist, A., & Hinojosa, T. (2001, July). Aligning GLOBE to national and international standards. Paper presented at the 6th Annual GLOBE Conference, Blaine, WA.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Thousand Oaks, CA: Sage.

Reiser, B. J., Spillane, J. P., Steinmuler, F., Sorsa, D., Carney, K., & Kyza, E. (2000). Investigating the mutual adaptation process in teachers' design of technology-infused curricula. In B. Fishman & S. O'Connor-Divelbiss (Eds.), 4th International Conference of the Learning Sciences. (pp. 342–349). Mahwah, NJ: Erlbaum.

Rowan, B. (2002). The ecology of school improvement: Notes on the school improvement industry in the United States. Journal of Educational Change, 3(3/4), 283–314.

Rowan, B., Camburn, E., & Correnti, R. (2004). Using teacher logs to measure the enacted curriculum: A study of literacy teaching in third-grade classrooms. Elementary School Journal, 105(1), 75–102.

Rowan, B., & Miller, R. J. (2007). Organizational strategies for promoting instructional change: Implementation dynamics in schools working with comprehensive school reform providers. American Educational Research Journal, 44(2), 252–297.

Sarama, J., Clements, D. H., & Henry, J. J. (1998). Network of influences in an implementation of a mathematics curriculum innovation. International Journal of Computers for Mathematical Learning, 3(2), 113–148.

Scribner, J. P., Sawyer, R. K., Watson, S., & Myers, V. L. (2007). Teacher teams and distributed leadership: A study of group discourse and collaboration. Educational Administration Quarterly, 43(1), 67–100.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15(2), 4–14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1–22.

Smith, L. K., & Southerland, S. A. (2007). Reforming practice or modifying reforms? Elementary teachers' response to the tools of reform. Journal of Research in Science Teaching, 44(3), 396–423.

Smith, M., & O'Day, J. A. (1991). Systemic school reform. In S. H. Fuhrman & B. Malen (Eds.), The politics of curriculum and testing (pp. 233–268). Bristol, PA: Falmer.

Songer, N. B., Lee, H.-S., & Kam, R. (2002). Technology-rich inquiry science in urban classrooms: What are the barriers to inquiry pedagogy? Journal of Research in Science Teaching, 39(2), 128–150.

Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. Journal of Curriculum Studies, 31, 143–175.

Spillane, J. P. (2004). Standards deviation: How schools misunderstand education policy. Cambridge, MA: Harvard University Press.

Spillane, J. P., & Burch, P. (2006). The institutional environment and instructional practice: Changing patterns of guidance and control in public education. In H.-D. Meyer & B. Rowan (Eds.), The new institutionalism in education (pp. 87–102). Albany: State University of New York Press.

Spillane, J. P., & Callahan, K. E. (2002). Implementing state standards for science education: What district policy makers make of the hoopla. Journal of Research in Science Teaching, 37(5), 201–225.

Spillane, J. P., Diamond, J. B., Walker, L. J., Halverson, R., & Jita, L. (2001). Urban school leadership for elementary science instruction: Identifying and activating resources in an undervalued school subject. Journal of Research in Science Teaching, 38(8), 918–940.

Spillane, J. P., & Jennings, N. E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. Teachers College Record, 98, 449–481.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. Review of Educational Research, 72, 387–431.

Supovitz, J. A., & Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. Journal of Research in Science Teaching, 37(2), 963–980.

Swanson, C. B., & Stevenson, D. L. (2002). Standards-based reform in practice: Evidence on state policy and classroom instruction from the NAEP state assessments. Educational Evaluation and Policy Analysis, 24(1), 1–27.

The GLOBE Program. (2005). The "Next Generation" GLOBE (NGG): Developing strategic partnerships to improve science education, enhance environmental awareness, and increase understanding of the Earth as a system. Retrieved December 5, 2005, from https://globe1.globe.gov/webarchs/all/NGG_WhitePaper_9_291.pdf

Uekawa, K., Aladjem, D. K., & Zhang, Y. (2006). Impact of comprehensive school reform on social capital and pedagogical change. Journal of Education for Students Placed at Risk (JESPAR), 11(3&4), 295–307.

Warren, L. L., & Muth, K. D. (1995). The impact of common planning time on middle grades students and teachers. Research in Middle Level Education Quarterly, 18(3), 41–58.