Fostering Innovation Implementation: 
Findings about Supporting Scale from GLOBE

Barry J. Fishman
University of Michigan, 610 E. University, Ann Arbor, MI 48109, fishman@umich.edu

William R. Penuel
SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025, william.penuel@sri.com

Ryoko Yamaguchi
Abt Associates, 4550 Montgomery Ave., Suite 800N., Bethesda, MD 20814 , ryoko_yamaguchi@abtassoc.com

Abstract: This study explores factors related to successful implementation of the GLOBE science education program through an examination of outcomes related to professional development (PD), technical support, and resource provision. We analyzed results from a survey of 454 teachers served by 28 different PD providers using a 2-level hierarchical linear model. Results support findings from earlier studies about the importance of teachers’ perceptions of coherence with their own and their district goals for student learning and the value of active learning strategies. At the same time, our models indicate that the correlates of effective PD vary for different outcomes, and that for program implementation equipment and technical support are critical factors. The results provide guidance for researchers and innovation developers seeking to better understand the conditions which foster implementation, scalability, and sustainability.

The past several years have seen a growing focus on the interrelated topics of implementation, scale, and sustainability among learning sciences researchers (Dede, in press; Fishman, Marx, Blumenfeld, Krajcik, & Soloway, 2004; Squire, MaKinster, Barnett, Leuhmann, & Barab, 2003). The reason for this interest is clear. If the carefully-crafted and researched innovations emerging from the learning sciences community are to have real-world impact on educational systems and learning, it is critical that we refine our understanding of what factors are most important for shaping the way(s) that our innovations are adopted, adapted, and implemented. Thus far, most efforts in support of making innovations more scalable and sustainable have focused on providing professional development for end users (teachers), providing resources such as computers, and providing support in the form of technical assistance (Blumenfeld, Fishman, Krajcik, Marx, & Soloway, 2000). But as projects evolve beyond the initial stages of funded research, developers struggle with the financial and practical implications of all three of these forms of support. This paper provides empirical evidence about the relative importance of professional development (PD), resource provision, and technical support in fostering implementation. These data have relevance for policy makers and researchers seeking to understand how best to foster the implementation, scaling, and sustainability of learning sciences innovations and programs.

Theoretical Framework and Literature Review

Blumenfeld, et al. (2000), use the metaphor of usability to define a 3-D space for considering the total capacity of an organization to successfully adopt an innovation. The three dimensions of this space are defined as capability, culture, and policy and management. Of these three, capability is the most mutable, especially when one’s focus is on school organizations. The dimensions of culture and policy and management are part of what has been described as the “grammar of schooling,” which may be changed, but only with great difficulty (Tyack & Tobin, 1994). Capability, however, describes the technical skills and abilities possessed by actors within the school organization. These skills can be improved through well-designed programs of PD, through providing task-specific support, and providing resources to supplement those normally available within the organization (Blumenfeld et al., 2000; Fishman, 2005). The question for research, and our central question in this paper, is what kinds and attributes of PD and support lead to the most effective outcomes in terms of successful program implementation?

Recent research explores the complex linkages between the design of PD, teachers’ learning, and classroom practice at multiple levels of scale (Borko, 2004). One of the most notable recent studies of PD effectiveness is by Garet, Porter, Desimone, Birman, and Yoon (2001), which is the first large-scale study to examine how core and structural components of PD in the Eisenhower Math and Science program led to self-reported changes in teachers’ knowledge and classroom practice. Specifically, Garet et al. (2001) found evidence supporting the value of the
following structural features of PD: reform-orientation (with reform-oriented activities such as teacher study groups being more effective than traditional PD settings such as workshops or college courses), duration, both in terms of time span and total contact hours, and the collective participation of teachers from the same school. The following core features were also found to contribute to enhanced knowledge and skills and changes in teaching practice: a focus on content knowledge, active or inquiry-oriented learning approaches in the PD, and a high level of coherence with other reform activities and standards in the teachers’ local school context (Desimone & Porter, 2002). These findings are consistent with an emerging consensus about high-quality PD (e.g., Hawley & Valli, 1999; Kennedy, 1999), but because of the unique nature of the Eisenhower PD program and their study design, Garet et al. were able to demonstrate how these factors related to one another in order to produce changes in teachers’ self-reported knowledge and practice.

The scope of the PD program has tremendous implications for the kinds of things researchers can learn about teacher learning (Borko, 2004), and an appreciation of scope is critical to understanding the findings from the Garet et al. (2001) study. By scope, we refer to the breadth of PD goals and activities pursued within a particular program or portfolio of projects. There are key differences in the research questions that can be addressed to PD programs that focus on a limited set of goals and in just a few kinds of settings, versus those that have adopted a variety of goals (e.g., increasing teachers’ content knowledge and improving their ability to teach diverse students) and activities designed to advance those goals. Large-scale studies of PD programs with a broad scope can yield general models of effective PD, but are not as helpful for determining issues related to the implementation of specific programs. In addition, the questions one is able to pose on a questionnaire targeting multiple activities about, for example, subject matter content or specific program outcomes are necessarily general, and therefore harder to use to determine success or failure of implementation of a specific program model beyond a rough approximation. In the Garet, et al. (2001) study, for instance, the range of potential uses of technology supported by PD activities under the Eisenhower program limited researchers to asking teachers to report on changes in the cognitive challenge of classroom activities, the use of technology, and the instructional methods employed broadly. As a result, it would be difficult for any of the sub-programs that comprised the Eisenhower program to use responses to these kinds of questions to determine if the changes are consistent with their own program goals. This would include being able to determine if technologies are being used appropriately or if the pedagogical changes are consistent with the goals of the PD.

The scope of programs examined in the Garet, et al. (2001) study is also a limitation when it comes to understanding how to design PD for specific learning sciences innovations. Because Eisenhower is a funding program that supports a broad variety of different PD programs, a study of it may reveal designs that in general lead to generally desirable outcomes, but it is difficult to use data from the Eisenhower study to make claims about the effectiveness of any of the particular PD activities that comprised the Eisenhower portfolio. It is also impossible to say what the real impact of that PD was on any particular innovation that was supported by Eisenhower PD. Thus there is an opportunity to apply the types of methods used in the Eisenhower study to more highly specified programs, where it is possible to know more about target outcomes. An example of a study with a more appropriate scope for examining program implementation is by Kanaya, Light, and Culp (2005), who set out to test the applicability of Garet et al.’s (2001) structural and core features model to a program designed to prepare teachers to apply a specific curricular planning model to integrate technology into their lessons. Using data from a survey, they examined what features of PD predicted two program-related outcomes: using new computer applications in the classroom and creating technology-rich lessons. Although the duration of teachers’ initial PD varied, duration did not predict the first outcome. However, teachers’ sense that the experience prepared them to integrate technology into their grade level or subject and support their particular students, as well as the perceived usefulness of the pedagogical components of the workshop, were significant predictors of teachers’ technology use with students. By contrast, all three variables – duration, teachers’ perception of relevance to their classroom context, and pedagogical strategies used in the activities – predicted how much teachers reported implementing new technology-rich lessons.
Methods
Research Questions
Our study sought to examine the correlates of effective PD by addressing the following research questions:

*How do different features of PD, access to materials, and support contribute to self-reported changes in teacher knowledge and practice in a single inquiry science education program?*

*How do these PD features, access to materials, and support contribute to program implementation?*

Context
Our study was conducted in the context of GLOBE, an international science and education program that has trained over 24,000 teachers since 1995 and reaches roughly 200,000 students in the U.S. each year (Penuel et al., 2002). In addition to being a program that exists at scale, there is evidence from student assessment data that the program has a positive effect on student learning (Coleman & Penuel, 2000; Means et al., 2000). Though there is variation in what might be deemed a complete implementation of GLOBE in any one school, the bottom line for participation is that students be engaged in gathering data according to predefined scientific protocols and that the data gathered be reported to the GLOBE web site for use by scientists and other participants. GLOBE implementation places intensive demands on schools in terms of resources, technical support, and knowledge. For instance, schools need access to some relatively common science resources, such as water quality kits, but they also might need specialized resources, such as an instrument shelter for use in gathering meteorological data. Schools need access to the Internet and must have computers in locations where students and teachers can readily use them to input data to the GLOBE web site. And teachers need knowledge of how to use the scientific protocols, what the purpose of the protocols are in relation to big ideas in science, and how those protocols relate to and can be integrated with local curriculum standards and activities. In GLOBE, PD providers called “partners” are expected to design PD experiences for teachers to prepare them to implement the program and to support them afterwards. Partners’ vary with respect to how they design PD and support teachers, allowing us to study the effects of this variation on outcomes such as program implementation.

Sample
Our study focused on the support and PD activities of 28 GLOBE partners (universities or regional science organization who coordinate local training and support of GLOBE teachers). 454 teachers were in our sample, or a median of 11 teachers per provider (with a range of from 2 to 141 per provider). We focused on teachers who had received PD within the past two years and PD providers who offered the broadest range of PD opportunities in terms of the factors that Garet, et al. (2001) explored.

Data Collection
We developed two questionnaires. One was for GLOBE partners, and it inquired as to the nature of their PD offerings, and was administered in summer 2004. The second was developed specifically for GLOBE teachers, and was administered in the spring and summer of 2005. This questionnaire was largely based on the instrument used by Garet, et al. (2001), though we added additional questions that were GLOBE-specific with respect to changes in knowledge and teaching, as well as new questions about technical support and resource provision, which had been used in earlier GLOBE teacher surveys. We piloted this instrument using a cognitive interview approach to determine that the questions were interpretable by GLOBE teachers in ways consistent with our intended meaning (Desimone & Carlson Le Floch, 2004). A third source of data to inform our study is the web-based GLOBE Data Archive, which contains information related to all data reporting by GLOBE teachers, providing us with an objective implementation measure that is more reliable than self-report.

Measures
Dependent Measures
We used five dependent measures in our study: 1) Teacher Change (in classroom practices, 6 items, six items, alpha = .84); 2) Knowledge of Pedagogy (6 items, alpha = .91); 3) Preparation for Student Inquiry (16 items, alpha = .94); 4) Data Reporting (binary variable); and 5) Protocol Use (binary variable). Measures 1 and 2 correspond to the outcomes of reported changes to knowledge and practice broadly and measures 3-5 correspond to program implementation as an outcome.
**Teacher-level Measures**

The eight teacher-level variables related to PD and support experiences of teachers were: 1) Collective Participation (2 items, total score from 0-2); 2) Reform-like PD (an index of 6 types of PD activities); 3) Traditional PD (an index of 4 types of PD activities); 4) Coherence (6 items, alpha =.86); 5) Time Span (of PD, an index from 0 to 24 months); 6) Barriers (an index of 10 items, scored from 0-20); 7) Equipment (a binary variable); and 8) Technology Support (a binary variable).

**Provider-level Measures**

Four variables were used to measure the PD designs of the providers: 1) Planning Support for Implementation (an index of 12 PD instructional approaches); 2) Focus on Student Inquiry (an index of 13 PD instructional approaches); 3) GLOBE Content (a factor of 9 items, alpha = .92); and 4) Total Hours of PD offered (an index).

**Analysis**

Because we have data at two levels (partners and teachers) that are nested, we used hierarchical linear modeling (HLM) to examine teacher PD experiences (level-1) and the PD context that the partners offered (level-2) (Raudenbush & Bryk, 2002).

**Findings**

Three themes emerged from our HLM analyses. The first centers on the importance of teachers having PD experiences that are coherent, ongoing, and consistent with their local school and district goals. Coherence was a significant factor in predicting both GLOBE implementation and changes in teacher knowledge and practice. Specifically, Coherence was a significant positive predictor for Teacher Change ($\beta = .28; \text{s.e.} = .05; t\text{-ratio} = 5.59$), Knowledge of Pedagogy ($\beta = .59; \text{s.e.} = .05; t\text{-ratio} = 11.22$), Preparation for Student Inquiry ($\beta = .51; \text{s.e.} = .04; t\text{-ratio} = 13.18$), and GLOBE Protocol Use, a key measure of successful implementation (OR = 1.87; CI = 1.45, 2.41). These results show that the consistency and coherence of the PD has a positive impact on teacher learning and practice, which is consistent with the Garet et al. (2001) findings. Teachers who participated in more Reform-like PD activities, such as participating in a teacher collaborative/network, working with mentors, immersion programs, or study groups, indicated greater Teacher Change in Practice ($\beta = .17; \text{s.e.} = .06; t\text{-ratio} = 2.94$) and Preparation for Student Inquiry ($\beta = .11; \text{s.e.} = .04; t\text{-ratio} = 2.34$). The finding that reform-types of PD have positive influences on knowledge and changes in practice is consistent with the Garet, et al. (2001) findings.

The second theme that emerged is the importance of the PD context offered by the partners. In Garet et al.’s (2001) study, they found that teacher perceptions of active learning, which we interpret in our study as Support for Implementation Planning had a positive impact on teacher learning. Similarly, we found that when partners provide more time to consider how they will implement the program in their classrooms as part of their PD, teachers were more likely to have more confidence in enacting inquiry-focused learning and GLOBE protocol use. Specifically, Support for Implementation Planning activities positively predicted Preparation for Student Inquiry ($\beta = .03; \text{s.e.} = .01; t\text{-ratio} = 2.92$) and Protocol Use (OR = 1.13; CI = 1.04, 1.22). These results suggest that the PD context mattered in teachers’ use of GLOBE protocols and enhanced knowledge and skill of science inquiry. Total Hours of PD provided by partners also mattered in whether teachers used the GLOBE protocols with students and knowledge of science inquiry. Interestingly, in both cases, there was a negative relationship between total hours of training and teachers’ enhanced Preparation for Student Inquiry ($\beta = -.01; \text{s.e.} = .00; t\text{-ratio} = 2.26$) and their implementation of the GLOBE scientific protocols (OR = 0.96; CI = 0.93, 0.98). For the most part, what mattered for GLOBE implementation was different from what mattered for changes in teachers’ knowledge (both pedagogical and science inquiry) and practice. First, it appears that the total amount of GLOBE-related scientific content emphasized during the PD had a significant positive relationship to teacher’s enhanced Preparation for Student Inquiry ($\beta = .17; \text{s.e.} = .05; t\text{-ratio} = 2.26$). Second, PD providers who had a greater focus on student inquiry and the scientific process had a significant positive relationship to teachers’ Data Reporting practices (OR = 1.23; CI = 1.06, 1.42), while focus on student inquiry was not significant for other outcomes, such as teacher change and knowledge. This suggests that an emphasis on the scientific process and inquiry in the PD was more important to faithful implementation (as measured by data reporting) than to more general changes to practice.
The third theme is that technical support and provision of resources are important to the success of program implementation. Receiving Technology Support had a positive impact on teachers’ Knowledge of Pedagogy ($\beta = .24$; s.e. = .10; t-ratio = 2.49), Data Reporting (OR = 2.45; CI = 1.35, 4.46), and GLOBE Protocol Use (OR = 2.40; CI = 1.28, 4.52), though it was not associated with Teacher Change or Preparation for Student Inquiry. Not surprisingly, being provided with Equipment had a positive impact on teachers’ Data Reporting behaviors (OR = 1.99; CI = 1.21, 3.28) and GLOBE Protocol Use (OR = 1.87; CI = 1.16, 3.01). Finally, the more Barriers teachers faced, such as not having enough time to prepare, the less likely they were to report data (OR = 0.90; CI = 0.83, 0.97).

**Discussion and Conclusion**

It is important that our overall findings about predictors of changes to knowledge and practice reinforce those of Garet, et al. (2001), because the challenge for those who design and provide PD is so critical. There is a tremendous demand for improved teacher quality, and a great ongoing need for continuing professional learning among practicing teachers at all points in their careers (Feiman-Nemser, 2001). It is necessary, therefore, for the field to develop valid and reliable guidance on how that PD is best structured to lead to various desired outcomes. Using HLM allowed us to incorporate information from PD providers directly into our models, and our examination of PD in a specific program context yielded important insights about what constitutes effective PD. This exploration revealed an interesting challenge for interpreting the linkage between effects on teachers’ knowledge and beliefs and related changes to classroom practice and program implementation. In short, what we learned is that even when PD leads to changes in knowledge, beliefs, and (in a limited sense) practice, that does not necessarily mean that these changes translate into more effective program implementation. The three factors in our study which arose as most important for supporting implementation are coherence, planning support for implementation, and technical support (both in terms of providing GLOBE equipment and giving assistance in using it). We will consider each below, and also in what ways these three factors may be pointing towards the same issue.

Garet, et al. (2001) focused on three dimensions of coherence: connections between the content of PD and other goals and activities teachers are engaged with, alignment with state and district standards and assessment frameworks, and the amount of communication with other teachers around the goals of the PD. In extending our thinking about coherence to include PD related to specific programs or curricula, we also draw on the work of Spillane and colleagues (Spillane, Reiser, & Reimer, 2002), who argued that teachers’ interpretations of policies and goals for learning influence their views of particular innovations, and related work that points to how teachers are the ultimate arbiters of the “meaning” of educational policy, and how school administrators’ priorities can often conflict with the original intent of policy (Spillane, 2004). Our own analyses concur that coherence is a significant variable at the teacher level, not just at the PD level, and is best understood by examining teachers’ perceptions of coherence. It seems critical to use that the construct of “coherence” be expanded to include a better understanding of how teachers see the connections between their own goals and those that lay behind PD. A first step may be the explication of those goals by PD providers, in order to make them more explicit to participating teachers. A second step involves eliciting teachers’ own goals for participation and their district’s goals for student learning, as well as the alignment (or misalignment) among curriculum materials, PD, and teachers’ and districts’ goals.

A second factor of interest is the significance of partners’ efforts to incorporate planning support for implementation into the design of their PD activities. In many cases, partners incorporated opportunities for hands-on learning and time to help teachers plan for implementation into what could be called traditional workshops that occurred away from teachers’ work sites. The finding that planning support—indepedent of whether the methods were incorporated into a formal training session or classroom-based support—is a significant factor in predicting the effectiveness of PD provides important lessons for providers about how to design “traditional” workshops so that they can be effective. Indeed, many teachers in GLOBE implemented the program despite participating in only traditional workshops; yet most of these workshops did incorporate extensive opportunities for teachers to practice skills needed for implementation and to plan to integrate the materials from GLOBE into their regular curriculum.

The third factor of interest in the present study is technical support, which includes the provision of materials, technology, and equipment necessary to successfully participate in a program. In the case of GLOBE, this included scientific data gathering instruments, access to computers, and help with using the GLOBE data-entry web pages. Our finding was clear—this factor was the most important in terms of whether teachers actually participated in a primary measure of success for the GLOBE program, the reporting of data. We argue that this is a key
consideration in the usability of the program or innovation. Teachers may alter their knowledge or beliefs, but they won’t act on these changes in the classroom (at least not in ways that the program developers might desire) because they will ultimately lack the technical capability to follow through (Fishman, 2005). Our finding is consistent with previous research on barriers to the success of curricular innovations, which found that it is critical to achieve a match between the capacity demands of any particular program and the available capacity in schools (Blumenfeld et al., 2000). Some of the capacity building can take place at the level of teachers through PD, but there needs to also be capacity building at the organizational level through avenues such as increased technical support to make a program work in individual classrooms. Support may also be related to coherence. It is likely that if a teacher does not feel that the program is receiving sufficient support, either from outside sources or within their district, they are less likely to make the effort required to fully participate.

In our study, not only the provision of equipment, but also help in setting it up was important. In the case of GLOBE, such help is often critical, since teachers must often assemble materials and follow specific guidelines for locating instrument shelters needed to collect data. Moreover, as our other case study research has shown, partners often combine visits to classrooms in which they provide more “reform-type” PD such as model teaching to teachers with troubleshooting equipment problems (Penuel, Shear, Korbak, & Sparrow, in press). Thus, equipment support might better be understood as integral to PD, not simply as a necessary but independent requirement for successful implementation.

We would be among the first to note that, when examined as discrete elements, what our data shows is not surprising. One would expect professional development that is coherent and consistent with teachers’ other work to be more effective, and one would also expect professional development focused on inquiry to be more effective in fostering inquiry-oriented teaching in the classroom. Nor is it surprising that teachers who receive technical support or are provided with equipment are more likely to successfully implement a program that demands teachers implement new kinds of technology in order to participate. What is new and important about these findings is looking at them in combination in order to understand the relative importance of professional development design and support, as revealed by the HLM model, in terms of fostering an innovation that can be implemented by teachers, and ultimately be sustained. There are few programs today operating at sufficient scale where we can examine these elements in juxtaposition, and data of this type are critical for informing the broader policy debate about how to fund and support research on educational innovations. From a learning sciences perspective, this study addresses a basic and recurring question for researchers: “What happens when I am done with the development phase for my innovation?” This also raises critical issues for those engaged in the funding of technology-rich programs in education, at both federal and local levels. As a community, the learning sciences only recently convinced policy makers that it was important to do more than simply purchase hardware and software for education – it is also critical to provide professional development to help teachers learn how to use new tools and integrate them into their instruction. But the data from this study suggests that there is an ongoing relationship between the two activities, and that continuing technical support is as important to program implementation as professional development.

References


Kennedy, M. (1999). *Form and substance in mathematics and science professional development* (NISE Brief No. 3(2)). Madison, WI: National Center for Improving Science Education.


**Acknowledgements**

This work was supported by National Science Foundation grant #0223068. We also thank Kenny Nguyen, Chris Korbak, and Amy Lewis for their support in conducting this research. The opinions expressed in this work are those of the authors and do not necessarily reflect the views of GLOBE, the NSF, the University of Michigan, SRI International, or Abt Associates.