Research and Practice in Technology Enhanced Learning
Vol. 2, No. 1 (2007) 51–74

DESIGNING FORMATIVE ASSESSMENT SOFTWARE WITH TEACHERS: AN ANALYSIS OF THE CO-DESIGN PROCESS

WILLIAM R. PENUEL*, JEREMY ROSCHELLE and NICOLE SHECHTMAN†
SRI International, Menlo Park, CA 94025, USA
**william.penuel@sri.com*
http://ctl.sri.com/

Researchers in the learning sciences have explored a collaborative approach to developing innovations that fit into real classroom contexts. The *co-design* process relies on teachers' ongoing involvement with the design of educational innovations, which typically employ technology as a critical support for practice. To date, investigators have described the application and results of co-design, but they have not defined the process nor explored how it plays out over time. In this paper, we define *co-design* as a highly-facilitated, team-based process in which teachers, researchers, and developers work together in defined roles to design an educational innovation, realize the design in one or more prototypes, and evaluate each prototype's significance for addressing a concrete educational need. We suggest seven key process components and use data from a systematic set of interviews to illustrate the roles of teachers and researchers in co-design and describe how tensions in the process can unfold and be resolved over time.

Keywords: Co-design; formative assessment; 1:1 mobile devices.

1. Introduction

We know from studies of educational innovations that teachers' interpretations of their classroom contexts strongly influence the level and nature of adoption of those innovations. For example, teachers who perceive their principal as supportive and their school as a collegial environment are more likely to adopt innovations (Penuel, Frank & Krause, 2006; Talbert & McLaughlin, 2002). In such environments, teachers are more likely to ask for and receive help they need to implement new classroom strategies (Rosenholtz, 1989; Rosenholtz & Simpson, 1990). Similarly, researchers have found that when developers of innovations are able to match their programs or curricula to teachers' goals for their own students' learning and to their district's requirements, teachers are more likely to implement those innovations (Cohen &

†Authorship of this article is co-equal. Names of the authors are in alphabetical order. Nothing should be inferred from the order of authorship as to differences in the level of authorship for this paper.

Hill, 2001; Fishman, Penuel & Yamaguchi, 2006; Garet, Porter, Desimone, Birman & Yoon, 2001).

Similarly, studies of technology-supported innovations in schools show that adoption depends in large measure on how well teachers perceive them to fit within their goals for students, teaching strategies, and expectations for student learning (Blumenfeld, Fishman, Krajcik, Marx & Soloway, 2000; Means, Penuel, Korbak *et al.*, 2001; Means, Penuel & Padilla, 2001). Perceived fit for technology-supported innovations is a function partly of teachers' current teaching practice (Cuban, 2001) and partly of their beliefs about student learning and what kinds of tasks their particular students are capable of accomplishing successfully (Becker & Anderson, 1998; Windschitl & Sahl, 2002). But fit is also a function of the local context, including the social and technical capacity of schools and districts to support implementation (Blumenfeld *et al.*, 2000) and alignment of an innovation with standards for student learning (Means *et al.*, 2001; Means, Penuel & Padilla, 2001).

Researchers in the learning sciences have explored a collaborative approach to developing innovations that fit into real classroom contexts. The *co-design process* relies on teachers' ongoing involvement with design of educational innovations, which typically involve technology as a critical support for improving teaching practice. Innovations resulting from the co-design process include a wide range of curriculum materials in school science (Edelson, Gordin & Pea, 1999; Reiser *et al.*, 2000; Shrader, Williams, Lachance-Whitcomb, Finn & Gomez, 2001) and mathematics (Hand, Underwood & Nielsen, 2001; Roschelle *et al.*, 1999). In addition, the co-design process has been used to develop assessment materials in science to help students monitor their progress in inquiry (Atkin, 2001; Edelson, 2002) and in reading and mathematics to help teachers adjust instruction (Shepard, 1997).

The co-design approach to the problem of "fit" differs from top-down approaches in that the process pays close attention to teachers' everyday work practices and their classroom contexts. Like all reformers, co-design practitioners pay attention to broad goals for learning and for improving large-scale systems. However, in contrast to some reform approaches, co-design looks at broad reforms through teachers' eyes. Its goal is the creation of innovations that help teachers meet reform goals. In contrast to approaches to reform in which teachers are simply expected to follow scripts for teaching (Atkin & Black, 2003; Sawyer, 2004), teachers are active participants in co-design and are viewed as professional contributors to reforms.

The co-design approach has close affinities with several other traditions of design in industry, from which it borrows key insights and assumptions. Most of these traditions are part of a so-called "semantic turn" in design (see Krippendorff, 2006), in that they share a concern with designing artifacts that have meaning and make sense for users in their context. For example, in its attention to teachers' goals for learning, co-design parallels the tradition of *value-sensitive design*, in which successful adoption is assumed to depend on whether designs reflect users' core values and needs (Friedman, 1996; Friedman, Kahn & Borning, 2003). Co-design requires attention to the usability of designs in particular learning contexts, which makes it similar in focus to *user-centered* (Carr, 1997) and *scenario-based* design processes (Carroll, 1995). In its commitment to the active participation of teachers in the design process, co-design shares many of the same values of *participatory design* (Ehn, 1992; Muller & Kuhn, 1993; Muller, Wildman & White, 1993), a tradition with a rich history in the development of workplace technologies. Finally, in its commitment to seeking frequent user input and working within the school year cycle to accomplish concrete goals, co-design is similar to processes used by designers who engage in *rapid prototyping with teams* and *team-based design* (Bowers & Pycock, 1994; Gorden & Bieman, 1995; Snyder, 2003; Tripp & Bichelmeyer, 1990).

Co-design also has specific parallels with traditions of design in education. In that the focus of most efforts is on developing tools that can enhance teaching and learning processes, it shares assumptions of researchers who have advocated for *learner-centered design* (Soloway, Guzdial & Hay, 1994), which emphasizes the need to develop tools that can motivate students in a wide variety of contexts. It also shares assumptions with researchers who conduct *design-based research* (Barab & Luehmann, 2003; Cobb, Confrey, diSessa, Lehrer & Schauble, 2003; Design-Based Research Collective, 2003; Soloway *et al.*, 1994) on the effects of innovations in education.

In this paper, we will define *co-design* and discuss the experiences of participants at the various stages of design. We will describe a case study of a specific co-design process we underwent to design software for handhelds that teachers could use to enhance their formative assessment of student learning. We will highlight the key process steps, as well as the dynamics and tensions from the perspectives of the participating researchers, software developers, and teachers. Awareness of these process steps, dynamics, and tensions can guide future co-design teams to establish conditions that are conducive to satisfying and productive co-design experiences.

2. Co-Design: A Definition

We define *co-design* as a highly facilitated, team-based process in which teachers, researchers, and developers work together in defined roles to design an educational innovation, realize the design in one or more prototypes, and evaluate each prototype's significance for addressing a concrete educational need. Although there are significant roles for teachers as participants in co-design, it is not a fully democratic process. Instead, accountability and ultimate responsibility for decision making rests with the project leaders, who are ultimately responsible for the quality of the educational resource being produced. In this section, we describe seven characteristic features of co-design as a method.

Co-design takes on a concrete, tangible innovation challenge. Co-design begins with a goal of creating some kind of innovation that seeks to advance an educational goal. In this respect, co-design is closely related to participatory design in that it is always oriented toward the development of a specific sociotechnical system — that is, a set of technologies and social activities in which the technologies are to be used — that meets a particular goal. At the same time, this aspect of co-design distinguishes it from open-ended, exploratory research on learning innovations in which the goal is to explore a problem space or theoretical construct. Success in co-design depends critically on whether a specific challenge — defined at the start — is met by the team of people working on the innovation.

The process begins by taking stock of current practice and classroom contexts. The co-design process begins with researchers' conducting fieldwork to understand how potential adopters of an innovation (teachers and students) perform tasks that are expected to be transformed by the innovation. Ideally, this fieldwork also documents key elements of the context — classroom and student characteristics, community values and perspectives, and local reform initiatives — that are likely to affect the enactment of the innovation. In documenting practice through fieldwork, "taking stock" in co-design has many similarities to rapid ethnographic techniques that are used in the design of workplace technologies (Hughes, King, Rodden & Andersen, 1995; Millen, 2000).

Co-design has a flexible curricular target. The process of assessing current practice inevitably requires co-designers to refine their design challenge and to construct metrics of success that are appropriate to teachers' contexts. In this respect, the curricular target must be somewhat flexible, allowing for a range of possible technical and social realizations. Flexibility is also important because the input of teacher participants in the design process is likely to alter researchers' conceptions of an educational resource. Co-design differs from traditional "waterfall" design in its responsiveness to users; it is related to rapid prototyping methods, which gather user input iteratively.

Co-design needs a bootstrapping event or process to catalyze the team's work. By "bootstrapping," we refer to the process by which a shared experience — such as a design workshop or retreat — can help build a common understanding of the need for a particular resource and a shared sense of what characteristics the resource must have to be of help to teachers in the classroom. If this shared experience happens away from the sites of practice for everyone — developers, researchers, and teachers — it can allow for focused attention to the task of catalyzing design teams' work.

Co-design is timed to fit the school cycle. Because it involves teachers, the work of design must be timed to fit the school calendar and teachers' work schedules. To involve teachers in the process, meetings must be scheduled toward the beginning or end of their school day or during common planning periods. Prototyping and testing need to happen during the regular school year, when students are present to help teams develop a realistic sense of what level of teacher effort will be required to enact the innovation's social and technological components. Extended design workshops are often scheduled for the summer, when teachers are free of school day pressures.

Strong facilitation with well-defined roles is a hallmark of co-design. Researchers facilitate by helping keep the team focused on the educational goals of the innovation; software developers help the team attend to issues related to feasibility and the development process. Both researchers and software developers actively work to keep everyone involved and excited through what is often perceived by teachers as a long process. For their part, teachers are expected to play active, if not equal, roles in the design process. Teachers help construct the key problems co-design must address, help frame the vision for what is to be created, test the innovation in their classrooms, and provide input as changes and refinements are made.

There is central accountability for the quality of the products of co-design. Unlike some approaches to participatory design, in which there is formal, shared accountability among workers, designers, and managers for system design, co-design has a central accountability structure for promoting quality in the co-design of educational innovations. Often, this accountability rests with the Principal Investigator of the grant that funds the effort. Central accountability means that those responsible for the design must be able to vouch for the quality of the resource created through the process. Deferring to users' judgments of quality is certainly a strategy that designers could choose, but ultimately in co-design, designers make the decisions about whether a resource is likely to benefit teachers and students.

3. The Dynamics of Co-Design

Co-design often elicits strong emotions among participants. The process of coming to terms with one's role and with the possibilities contained within the scope of a single project often brings alternating feelings of joy and frustration, excitement and boredom. Past research reveals tensions that occur in co-design. We review these briefly below and then examine data to illustrate how these tensions played out dynamically over time in a particular instance of co-design.

3.1. *Tensions revealed in prior studies*

From studies of participatory design we can learn about the tensions associated with engaging users in that process. For example, studies have pointed out that end users and designers often have different criteria for evaluating success (Blomberg & Henderson, 1990). These differences arise because users and designers occupy different professional worlds and thus attach different goals to the task of developing software. For example, overly technical discussions among software developers exacerbate barriers between developers and users (Bowers & Pycock, 1994). Success or failure, participatory design researchers point out, often turns on whether there is mutual understanding of the unique competencies and contributions of different team members (Bowers & Pycock, 1994). Educational applications of participatory and co-design ascribe similar importance to developing a common understanding of goals, roles, and contributions of team members (Shrader *et al.*, 2001; Shumar, 2003). At the outset of the project, teachers' notions about curriculum and about goals of projects often differ from developers' notions (Shrader *et al.*, 2001). Teachers often see researchers' solutions as too theoretical and not practical enough for real classrooms (Shrader *et al.*, 2001). By contrast, researchers often view teachers' limited content knowledge as a barrier to their contributing effectively to design efforts (Brown & Edelson, 1998).

Certain differences between workplace norms of teachers and researchers pose a second challenge to the process (Reiser *et al.*, 2000). Through the course of the co-design process, teachers' concerns often focus on usability and on students' capabilities much more than on teaching considerations. By contrast, researchers tend to take a more analytical stance and push teachers to justify goals. Teachers can find such talk unsettling, because it violates norms of interaction with colleagues that emphasize informal sharing and avoidance of peer critique (see Wilson, Miller & Yerkes, 1993). Often, therefore, co-design requires designers and users to engage for an extended time in bridging gaps by creating a common language to describe users' contexts (Adler & Winograd, 1992; Crabtree, 2003; Friedman *et al.*, 2003; Rheinfrank *et al.*, 1992). Here, the central challenge is to describe users' contexts in a way that gets beyond abstract characterizations of activities to provide concrete details designers can use as anchors for developing system requirements (Crabtree, 2003).

Time and productivity pressures are yet another threat to the success of co-design. A principal complaint about co-design is that it is very time intensive. In practice, involving users in design is complicated; it takes time, and users' input is often contradictory (Rheinfrank *et al.*, 1992). Many users find the design process boring (Ehn, 1992). Designers must sift through the contradictory input; furthermore, they are often the only ones with the expertise to implement a new program or software design. Attending to these difficulties is a central problem of co-design.

3.2. *Dynamic patterns over time: Trajectory of researchers' and teachers' experiences of co-design in one project*

In the current study, we sought to investigate participants' experiences and roles in co-design as they evolved in one project. Appreciating how the dynamics unfold over time can help researchers and teachers understand how the tensions of co-design can be addressed in the context of a project. Further, this knowledge can help participants in the co-design process recognize what kinds of confusions and difficulties are possible and perhaps likely at different phases of a project. We present researchers' and teachers' accounts of the process here to illustrate particular strategies co-designers may adopt for addressing tensions, confusions, and difficulties that may arise. We recognize that each project necessarily will have its own challenges and must chart its own course through those challenges. However, we hope by presenting one project's experiences we can provide strategies that could be adapted by other co-design teams.

The particular design challenge that is the focus of this paper was to explore how handheld computers might support improved classroom assessment in science classrooms at the middle grades. Project WHIRL, funded by a 3-year (January 2003 to December 2005) research and development grant from the National Science Foundation, was a partnership between SRI International and a medium-size district in the southeastern United States. Each school year of the project corresponded to a distinct phase of development. In the first year, we employed rapid ethnographic techniques to understand teachers and their work contexts (see Penuel, Tatar & Roschelle, 2004, for findings of this phase of research). In the second year, we engaged a small group of teachers in the process of co-design of new handheld software and in testing this software in their classrooms. In the third year, we recruited additional teachers from the district to participate in a field test of the software and provided them with professional development to enable them to use the software effectively in their classrooms. Results of the field test are presented in Yarnall, Schechtman and Penuel (2006).

Our choice of formative assessment as the focus of our design challenge was grounded in research that has documented powerful effects of improving classroom assessment as a strategy for advancing students' understanding of and abilities for inquiry (Black & Harrison, 2001; National Research Council, 2001; Petrosino, 1998; White & Frederiksen, 1998). Handheld computers, we believed, had the potential to overcome limitations of traditional assessments, namely that teachers find it difficult to collect and manage assessment data on a regular basis (see, e.g. Darling-Hammond, Ancess & Falk, 1995). We believed that the portability of handheld computers, as well as their capacity for storing and aggregating data and allowing for anonymous responses, would enable more frequent assessment by teachers and enable students themselves to become more actively involved in self-assessment. The low cost of the handheld devices themselves would also make the intervention plausible at scale.

A key tension in the design process for us, and one that motivated our case study, was a tension that emerged as researchers sought to instantiate a particular vision of formative assessment in the software that could be implemented by teachers. The research team had a commitment to involving students actively in self-assessment in ways that would be consistent with research on how people learn (e.g., National Research Council, 1999), what is alternately called *student-centered classroom assessment* (Stiggins, 1997) or *assessment for learning* (Shepard, 2000). Yet teachers tended to view the opportunity to improve assessment as an opportunity to become more efficient at giving traditional tests; that is, they wanted to become better at *assessment of learning*. The difference in viewpoint was important to us as researchers to address, since studies on formative assessment show that assessment *for* learning, but not necessarily assessment *of* learning, can contribute to improved learning outcomes (Black & Wiliam, 1998). Here we describe how we resolved this tension in the context of our own project, and how the resolution and inevitable compromises were experienced by the members of the co-design team.

4. Methods

4.1. *Study participants*

The teacher participants in the co-design process were all teachers of students in grades 4–9 in a mid-size district in the southeastern United States. The teachers were selected through an application process. Teachers who were interested completed a two-page application containing questions about their backgrounds, experience with technology, and approaches to instruction. No specific requirements for technology proficiency were designated. The goal of the application process was to recruit a group of teachers that was diverse with respect to prior experience with technology and approaches to teaching. Ten teachers applied, and all were accepted by a joint team of SRI researchers and district staff who reviewed the applications. Only seven were able to attend the week-long design conference to kick off the project.

The teachers who participated in this project were five women and two men, and they were an experienced group ($M = 18.5$ years of teaching science, $SD = 9.0$, range = 9–34 years). They were all white, but the classrooms they taught were diverse, with some teaching in classrooms where the majority of students were African-American or eligible for free or reduced-price lunches. Despite their teaching experience, only three had ever used a used a handheld computer before and only two used technology with their students more than once per week. We also employed a liaison who provided technical support locally to teachers. The technical support person was a former high school science teacher who participated in the design process and frequently helped teachers with both technology and instruction.

The SRI participants in the co-design process were themselves a diverse group. They included one psychologist, two cognitive scientists, one mathematician, two software engineers, and a former science teacher who specialized in assessment design. Unlike the teachers, all the SRI participants were experienced in using technology, and all but one had participated in or even led design efforts for educational technology applications in mathematics or science. Two team members had extensive experience in leading participatory design projects as well. The SRI participants were six men and two women, and five of the eight members were white. Two members were Asian or Asian-American, and one was Hispanic.

4.2. *Sources of data*

A primary source of data on participants' roles and experiences was a collection of interviews, conducted roughly once per month with each of the team members (including researchers and software developers) throughout the design process. In order to help teachers feel comfortable sharing their experiences, the interviews were conducted one-on-one in private by a researcher not involved in any of the co-design teams. Because the interviewer was located in California, the interviews with the teachers were conducted by telephone. While this form of communication can be quite distant, the researcher developed relationships with the teachers over the course of the year, through both monthly phone conversations and on-site classroom observations of some teachers. Interviews took place from September 2002 to May 2003, covering the full course of the design process. The structured protocol asked about the participants' contacts and activities, how the participants saw their role on the team, what decisions or progress had been made, what was going well, and what was not going well or was a concern.

In addition to interviews conducted during the co-design year, we relied on minutes of project meetings, researchers' interpretations of interviews and observations conducted with teachers during the first year of the project, and project artifacts used to help guide the design process. We used these secondary sources to help contextualize researchers' and teachers' comments in terms of the project's design activities.

4.3. *Approach to data analysis*

With this corpus of qualitative data, three researchers, two who had been part of the design process and the one who had conducted the interviews, used a grounded theory approach (Glaser & Strauss, 1967; Strauss & Corbin, 1990). With this approach, we aimed to develop a fresh theoretical interpretation of what the events of the co-design process meant to each of its participants. We did two phases of coding. In the initial phase, we sought to understand the data by attending to the general context, participants and their roles, timing and structuring of events, and the relative emphasis participants placed on various issues. In the more selective and conceptual second phase, we focused our interpretations on coding the central emergent themes of the evolution of roles and experiences of all the participants. We then used these codes to structure short memos articulating the major events, that would then be sorted into a coherent descriptive narrative. We note that this analysis is subjective and could possibly have rendered different interpretations. We consider that the themes we have drawn out here are to be taken as illustrative of how tensions unfold and are resolved over time in the context of co-design, helping us to refine our theory of co-design rather than pointing to empirical evidence of its overall success as a method for developing technology-supported innovations in schools.

Ours is necessarily a *researchers'* synthesis of participant experiences. We have not sought to validate our interpretations of teachers' experiences with the teachers themselves, nor have we sought their perceptions of what the researchers must have experienced in the co-design process. As such, we believe our work will be of most value to other researchers and not necessarily of immediate value to practitioners seeking to understand what it might be like to participate on a co-design team.

5. Overview of the Design Process for Project WHIRL

Although our design process was iterative in nature, we did set out to complete an initial, usable suite of software applications within one school year. Broadly speaking, there was a marked start-up phase or beginning phase of the design, in which design teams were formed and created charters and scenarios during summer 2002. The bulk of the 2002–03 school year was spent refining scenarios, developing requirements, and constructing and testing prototypes for the software. Toward the end of the school year, software developers created and tested versions of the software that would be used in the subsequent school year by 18 teachers in their

Table 1. Major goals and tasks of the project WHIRL co-design process.

Phase	Researchers' Goals	Major Tasks
Beginning	Create a shared vision for how software applications can support formative assessment	Research on Current Practice Design Conference and Team Formation Develop a Team Charter Develop Scenarios of Use
Middle	Articulate how students and teachers will interact with the software and in the classroom	Refine Scenarios of Use Develop Requirements Construct Paper and Technology Prototypes Test Prototypes in the Classroom
End	Develop applications that are usable by a broad range of teachers	Develop software applications Test software applications internally and in teachers' classrooms Create supporting documents and designs for teacher training

classrooms as part of a field trial. Table 1 describes the researchers' goals and major tasks for each phase of the co-design process.

We present Table 1 as background for helping readers interpret the perspectives of participants in the co-design process, which is the central focus of this paper. Our experience suggests that knowing the goals researchers had for particular phases of work may be helpful in this respect, but it is by no means sufficient for understanding the ups and downs of the process. Teachers' perspectives on the design process are especially important to consider, since in our case they did initially share the same map or birds-eye view of the design process; their understanding of what it means to be involved in co-design evolved over time.

6. Researchers' Experiences of the Co-Design Process

6.1. *The beginning: Making new software to support old practices?*

By the beginning of the co-design process with teachers, based on the initial ethnography in the first year of the project, researchers had already formed several strong impressions of the district and its teachers. Three of the research team members had conducted extensive interviews with teachers (both those on the co-design team and others) and observed several classrooms as teachers led instructional and assessment activities. The team members involved in the research became concerned about the fact that we did not observe much evidence of student-centered, learning-sciences-based teaching and assessment in science anywhere in the district. We had anticipated developing software that would support the more student-centered teaching we had assumed was in place from our initial visits to the district. Instead, we found practices that might be called "transitional," in that they included a mix of traditional, didactic teaching with hands-on, cooperative activities (Roschelle, Penuel, Yarnall & Tatar, 2004). Despite our concerns, some members of our team felt it would be particularly important to be more aware of our own biases on the co-design process. We suspected from our own internal discussions that it would be challenging not to push too hard to realize all the different goals we as researchers had for creating more opportunities for student-centered assessment. At the same time, we agreed that we would seek to build on those practices that indicated that teachers were in transition toward more student-centered modes of instruction.

In addition, the research team learned informally through conversations with teachers that the district's reputation for educational technology innovation had been based largely on administrator rather than teacher efforts. Teachers often reported being left out of the decision-making process, and many felt that they were ill prepared to implement initiatives that the administrators imposed on them. This situation struck us as potentially threatening to our own project, because we had developed our partnership initially with district administrators rather than with teachers. Also, the practices we observed in the first-year ethnography were from a broader sample of teachers and not necessarily the practices the co-design teachers would want to improve. When we shifted from being outside observers to inside co-designers, there was a need to gather additional data on the practices of our co-designers. We recognized the need to tailor the co-design process in such a way that teachers' voices could be heard first and given authority, even if administrators were somehow involved in the process.

We pursued several different strategies to address these sets of concerns. To attempt to balance the need for developing software applications that would support more student-centered forms of assessment with the classroom realities we faced, we decided on a structure that would leave ultimate accountability for applications with the research team but developed processes and artifacts designed to ensure teacher voice at each step of the design process. For example, we developed a set of internal documents to outline roles for researchers and software developers and establish milestones to monitor progress. One such document was a "Core Questions and Commitments" document to help guide district administrators' and researchers' participation in the design process. The "Core Questions" portion outlined our research questions and encouraged project leadership members to share these questions openly with participating teachers. The questions, moreover, were to be respected by all as questions for which we did not yet have good answers. We would encourage all teachers to see themselves as co-inquirers with researchers into these questions. Certain elements of our core commitments were also to be shared directly with teachers:

- To be responsive to teachers' concerns and those of other members of the school community (including parents and community members, as well as school leaders) in an ongoing way.
- To engage in and learn from a process of co-design of software.
- To learn from the use of technology and activities the role that handhelds can play in changing the focus of students' attention in the classroom toward their own thinking.

We planned to share these at the outset of the project, emphasizing that these commitments were part of our ongoing research programs and were built into the design of the project. We decided to share additional commitments that were based on our expertise — namely, our expertise as software developers, assessment experts, and evaluation researchers — in a way that respected practitioners' own perspectives on the topics in which we believed we had expertise. To that end, we decided on specific formats and activity structures within which we shared this expertise:

- By reflecting on contrasting "cases" of science teaching drawn from outside the district to help clarify that each member of the SRI design process also brought a unique perspective to the problems of teaching and assessment.
- By fostering and supporting teacher inquiry on documents or frameworks that were meaningful to us, through readings and discussion.
- By posing questions to teachers about consequences of particular design decisions that we might see, based on our prior encounters with similar types of classroom situations.

At our initial design conference to kick off the co-design process, attended by the members of the research team, the teachers, and some district administrators, we organized teachers and researchers into three design teams: Data Doers, Image Makers, and Hot-Q. A researcher led and facilitated each design team. Administrators were not part of any individual team; their role was to support the teams but not to be developers of ideas for each team. In subsequent teleconferences with design team members, administrators would not participate at all; however, a single district staff member charged specifically with teacher support participated in all calls for all design teams.

The initial design conference was only partially successful in helping the teams to develop a shared vision for software applications that could advance the goals of creating more opportunities for student-centered assessment in teachers' classrooms. Two of the teams formed did focus on supporting improvements to teachers' "transitional" practice, specifically the use of small groups for organizing laboratory science activities. The Data Doers team charter set out to help provide students with more immediate feedback when data they collected and recorded were implausible, "out of range," or in error. The software would function as a teacher's assistant in the classroom, prompting more immediate reflection on students' part about possible sources of error in measurement. A second team, the Image Makers, sought to enhance the use of drawing as an assessment mode by making enhancements to an existing software application called *Sketchy* developed at the University of Michigan. A third team, the Hot-Q team, however, was primarily interested in creating a flashcard-like application for the handheld that would enable the administration of traditional multiple-choice tests of vocabulary items on the handhelds. For this team, the gap between researchers' ideas and teachers' ideas about what to support was so wide that the researchers feared this team would not be able to design an application the researchers would consider to be of high quality.

As the design process got under way during the early part of the school year, researchers also noticed some hesitancy on the part of teachers to voice their concerns with researchers in teleconference meetings in which the group discussed different design scenarios. Especially on the Hot-Q team, some teachers were shy, and one researcher charged with facilitating one of the three design teams expressed being nervous about her ability to draw these teachers out and to encourage them to say "what's really on their minds." Teachers and researchers' second-guessed one another with respect to their opinions about particular designs, making it hard for the team to advance.

The software developers initially were quite frustrated as well, but for a different reason. They found that it was difficult during regular teleconferences with teachers to elicit enough details about teachers' classroom practice to develop requirements that would be adequate to teachers' needs. In addition, they struggled with what they heard to be competing visions for the software from teachers on the same team, and they sought, as part of teleconferences, to establish consensus as to what the software applications ought to be able to do. One developer noted in October, referring to his work on the Data Doers team:

> I had to clench my teeth... hearing for the 20th time [from a teacher], "It's really simple, if it's a ridiculous answer..." Getting teachers to consider the possibility that it's not quite as straightforward as they think it is, to see that it has more depth to it... I don't think we're trying to be pedantic researchers... if you are going to put it into software, you have to be specific. It's a learning process. We'll get there.

Two other research team members commented that the difficulties arose because teachers and researchers live in fundamentally different worlds. As one of them put it:

> Teachers live in a world where they have real students that they care about, and they can and, perhaps, have to be idealistic about knowledge and science. Researchers live in a world where design or data has to come together into a sensible, finished product.

Researchers facilitating all the teams remained concerned about the quality of the work they would ultimately produce. One researcher commented that he felt teachers had very low expectations of students' capabilities, and the teachers expressed anxiety about how the researchers would handle the situation. The researcher commented, "Everyone is concerned about worth of the projects. Are we going to be able to do more than drill-and-kill? We run the risk of making new software to support old practices."

At the same time, researchers were heartened by teachers' willingness to stick with the project — despite the lack of obvious progress toward developing system requirements — but were also keenly aware that they needed to provide concrete, tangible benefits to teachers for participation before too much time passed. A critical transition came when researchers visited teachers in early fall 2002 as a group, and each team presented on the progress they had made so far. In addition, at this meeting, one of the facilitators shared a road map of the design process for teachers, helping to clarify what researchers were increasingly perceiving to be teachers' limited understanding of the process. The meeting was a success, from both researchers' and teachers' points of view, giving the teams a sense that they had made tangible progress and a sense that they knew what would be happening next.

6.2. *The middle: Seeding ideas and bridging gaps*

In the middle of the design process (fall-winter 2002–03), the SRI team members' roles became much more differentiated. Researchers became much more focused on maintaining the forward direction of the team, and software developers began their work developing prototypes of the software for teachers to use. Researchers' roles as facilitators were becoming much clearer to them, and they also helped maintain the focus of teams on constructing an innovation that would meet the original needs identified in the teams' charter documents. The software developers played a more active role in team meetings at this point, sharing screen mockups and explaining how software might work to the teachers on the team. As the design process progressed, one software developer visited classrooms to see how the prototypes were working.

It became apparent that to meet the deadline of developing software that teachers could use with their students by the end of the 2002–03 school year, facilitators would need to take a more active role in moving teams forward. They were each surprised by how much effort it took to advance the team goals. As one put it, "I found that I have to play more of a directive role than I had thought." Another found the work of interpreting software developers' concerns to teachers exasperating at times, but felt that this role was especially critical to the success of the teams. The third facilitator found that active facilitation was necessary to coordinate the different work activities of the team, which included developing sample assessment activities, software development, and scenario refinement.

The more active roles played by facilitators were critical in helping resolve tensions researchers felt about creating software that could support innovative assessment practice. The facilitator on the Hot-Q team seeded the idea with teachers on that team that perhaps they could provide students with answers and students could pose questions about those answers. Although initially teachers were reluctant to pursue this idea, when the researchers agreed to find a software program already commercially available that could be used to quiz students, they found that when they tried having students pose questions as part of an assessment activity, the teachers learned a lot from the questions. As a consequence, the researchers grew much more confident that they were going to produce software that could support more inquiry-oriented teaching in science.

For their part, the software developers on the teams all became more active in pushing teams to refine and finalize specific requirements for software. In the background, too, they were identifying existing software applications and other tools that could support the teams' goals and facilitate the development process. As a result, the developers decided to form an alliance with two commercial software developers to develop refinements to existing applications, rather than build new applications from scratch for two of the teams. That decision allowed the team that had initially proposed developing a set of flashcards to develop software to promote student questioning. A third team developed its application using Satellite Forms, a tool for Palm OS designed to help build forms that can be used on handheld devices. Further, classroom visits by one of the developers helped bridge gaps between teachers' and software developers' understandings of classroom realities.

6.3. *The end: Getting to tools teachers can use*

Toward the end of the co-design process (spring 2003), developers and researchers were feeling both pressured by the limited time they had to prepare the software and energized that the co-design process had yielded an entire suite of software applications that excited the design team teachers. The roles of researchers and software developers remained differentiated, but the researchers' attention began to turn toward the field trial and the design of a research study to investigate how the handheld software would be used in the coming year. Software developers at that point were engaged in development and testing; they sought teachers' input with each release of the software and made adjustments to respond to the feedback. While the researchers were preparing to train new teachers in the use of the software as part of the field trial, software developers began quality assurance and other internal tests to make sure that the software would be robust enough to handle use by approximately 1,500 different students on more than 300 handheld devices.

The limited time frame for completing designs caused concerns among both the facilitators of the teams and the software developers. At the end of the process, facilitators were optimistic but felt that more work still needed to be done during the field trial. As one facilitator put it, "It still feels a little bit to me, particularly after interviewing these next-year teachers, we are producing a resource, just one of many, that these teachers can use. We're producing a type of paintbrush. They're the artist and they're doing the painting." Another commented, "We still need to iterate on it, and we need more time and resources to make it what it ought to be." Developers, for their part, noted that not all of the software features could be incorporated into the designs, because not enough time was available to resolve some technical problems. One commented that the addition of new features up through the spring posed problems as well:

> I'm a little concerned about some of the larger changes we're making to some of the software given the time period we need to test the applications and install them on all the devices for this coming year. I worry a bit

> about whether they've been adequately tested or not. There's maybe some fundamental tension between having the design teams producing what is essentially a series of prototypes with the idea that with each prototype you can introduce new features. But at the same time, if we're thinking about installing these on 13 sets of handhelds for the new year, we need to make sure they're adequately tested and work properly. There's always going to be a tension between adding new features at this stage of the game and having software that works reliably.

Like the facilitators, though, the software developers were optimistic, because they saw the eagerness and readiness of teachers to use the software with their students. As one put it, "That's a sign that there is a tool that they feel they can use."

7. Teachers' Experiences of the Co-Design Process

7.1. *The beginning: "I'm impressed that the team is sticking to it"*

In the first phase, teachers experienced many demands on their time and little that they could directly see or use. One of the teachers on the Image Makers team, for example, said in September 2002, "I feel like I'm putting a lot of effort into this but not getting anything in return. I haven't seen any fruits of my labor. I'm not sure they realize how much work is involved for the teachers." An essential dynamic of this phase was around building commitment to a long-term design process that was unfamiliar to teachers. The respect that teachers experienced was core to their continued involvement.

Teachers participated in several intensive interviews about their practices and shared information about the different types of lessons they did, the lab activities their classes did, and the problems they experienced in the classroom. They also provided their opinions about what their needs were and how tools might be able to help. The researchers had the teachers walk through detailed scenarios of how they imagined particular activities playing out in their classrooms.

Teachers on all three design teams felt good about their contributions. One teacher on the Data Doers team felt as though she was playing a critical role in helping the team understand the viewpoint of the teacher. Another teacher on that team stated, "They're sincerely interested in our opinions on things. We're the teachers. We let them know what works in the classroom, what needs to be tweaked, what a fourth or fifth grader is able to do, and how they handle these things." A teacher from the Hot-Q team, in December 2002, stated, "The fact that the researchers could incorporate our ideas into the software is outstanding. Usually in teams like this they don't take laypeople's opinions seriously. They were concerned about our concerns."

Teachers also voiced many concerns and frustrations during this phase. A central concern, which came up month after month for every teacher, was finding time in their already overbooked schedules to do all the work necessary for the design process. As one described, "You're 24 hours, 7 days a week working for and with your students. On my weekend, my personal time, that's the only time I get to work on this... It's super demanding to do all our responsibilities and serve on all our committees. It doesn't give us a lot of opportunity to play with technology. As anywhere, everything is due yesterday."

7.2. *The middle: "We're trying it out in the real world"*

As the projects moved forward into rapid prototyping (winter 2002–03), the teachers' roles became more multifaceted: they participated both as designers and as practitioners.

As designers, they described their role as "giving feedback," "input on what I see as potential for using the Palm," "tester," and "guinea pig." They were on the front lines trying out prototypes in their classrooms. As a teacher from the Data Doers team described, "We're trying it out in the real-world situation to see if what we think will work in theory will actually work in practice." They were observed by the researchers and developers, collected data about the students, and communicated to the team what was going on in their classrooms. They also worked on creating lesson plans to incorporate the tools, offered informed opinions about the software, and worked on more and more refined articulation of their needs. As a teacher from the Hot-Q team described, "One of my contributions is that I'm always looking for opportunities and ways to test the program out."

Teachers reported increased engagement during this creative phase. Many were putting substantial effort into finding new ways to integrate the prototypes or actual software into upcoming lessons. Several of the teachers described the excitement of their students with participating in the project. Many also discussed their appreciation of the researchers' openness and ongoing support, as well as their increasing coherence as teams.

Teachers also expressed several concerns and frustrations throughout this phase. Again, time was the most prominent issue. Teachers had difficulty finding the extra time to do all the planning and preparation for the activities.

In parallel with their work as designers, the teachers were also shifting and expanding their teaching practices. Several teachers reported growth over the year in how they were approaching teaching. They reported that the detailed reflection for both the design work and the issues around formative assessment expanded their awareness of their practices and increased their ability to reflect on how they know what students know. For example, a Hot-Q teacher team member reported as early as September 2002, "The process is pretty amazing to us. We are learning so many different ways that our problems can be solved." In February, she said,

> The people I'm connecting with are giving me a diversity of opinions to work with about best practices and worst practices. It's catalyzing my growth as a teacher and also myself as a person. It gives me a higher level of satisfaction and enthusiasm. For going into the classroom and teaching

> with my students, I know that there's something new and better for me to be doing, and a way of doing things more effectively. So it yields a more positive outcome in their learning. If you want to teach, that's everything.

7.3. *The end: "We're where we wanted to be — even if it was a bit slow"*

In the ending months of the co-design process (spring 2003), the teachers' roles shifted yet again. As a teacher from the Data Doers team put it, "I see my role as somewhat different than in the first and second years. The first year was development, and we're pretty much done with that. Now it's implementing." As rapid prototyping phased into software solidification, teachers could shift their focus even more toward practice. They spent more time thinking about how to integrate the handhelds into their classrooms.

Furthermore, the teachers found themselves in roles they described as "liaison," "salesperson," and "mentor." Two teachers on the Data Doers team found themselves talking with parents in the community about the project and the use of handhelds. One of them also felt that she was now trying to "sell the benefits of Data Doers to new teachers, especially the new members of the team." One teacher also had begun looking ahead to working with the new teachers next year: "It will be a mentor kind of thing." Another teacher also stated, "I'm anxious to see what happens when we add the new people. There are people on our staff who are excited. We'll build the team larger."

In the end, the teachers were pleased with the outcomes of the project, and all expressed satisfaction. A Hot-Q teacher, for example, stated, "I'm very happy about the project; to have had input and really seen it come to fruition is really nice. Good stuff." And a high school teacher from the Data Doers team, who had often expressed frustration over the course of the year, stated, "What's going well is that we got to the point at the end of the year that we wanted... I was frustrated it took so long. But I'm optimistic that we got done what we want to accomplish. We're where we wanted to be — even if it was a bit slow."

8. Discussion

Our analysis revealed that the key tensions of co-design persist across phases, yet the social dynamic between teachers and researchers evolves across phases. Persistent tensions include the following: teachers never have enough time; software developers and teachers tend to have different workplace norms; a common language is always a work in progress. Within these tensions, our analysis revealed an evolving social dynamic between teachers and researchers, particularly with regard to agency and ownership. In the beginning, teachers did not begin with a strong sense of ownership in the project or a clear sense of the roles that they would play in design. Our interviews with teachers revealed confusion early on about the goals of the project and about the design process itself. It was not until handheld computers arrived in their classrooms and classroom testing began that teachers began to feel as if the project was at least partly theirs. As teachers came to see how excited their students were about the handheld computers and how pilot versions of the software worked in their classroom, they began to adopt more active roles within the project. By the end of the 2002–03 school year, teachers had become strong advocates for the software. They recognized that they would play an important role in introducing and encouraging the use of the software to teachers who would participate in the field trial. But in some cases, teachers also advocated for the software in the broader school community.

Our analysis highlighted the role of SRI researchers as important in addressing the tensions and managing the positive evolution of the social dynamic and the outcomes. The researchers were those held accountable for the ultimate products of design, because the researchers had a fiduciary and technical responsibility to the funder, the National Science Foundation. Delegating this responsibility in the context of a federal grant would have required special permission from NSF, especially since we did not propose such a delegation of responsibility and because the school district was a subcontractor, and not the Principal Investigator. The accountability gave researchers some authority to help push teachers toward more innovative solutions; had the teachers or school district had authority for decisions, this might not have been possible. Instead, district priorities and initiatives might have shifted the goals and direction of the project in more fundamental ways. Thus, even though the process was participatory and tensions were addressed in ways that were intended to respect the wishes and ideas of teachers, researchers did not give up their vision for creating software that could lead to more student-centered assessment.

One important limitation of our study is that it does not allow us to contrast the quality of student learning resulting from co-design of innovations with innovations produced by other methods. Therefore, we cannot make strong claims that co-design produces better innovations or that those innovations are more likely to increase student learning. Future studies seeking to overcome this limitation might pose the question, “How does participation in co-design relate to better instruction and assessment practice among participants?” Indeed, answers to that question could provide useful insight into how taking on active roles within the design process might lead teachers to make deeper, more lasting changes to practice in ways that are supported by easy-to-use technology tools. We suspect that these changes in practice are essential mediators of any student learning gains resulting from an innovative tool design. Thus, researchers could begin to trace the pathway from co-design to changes in teacher practice and finally to student learning outcomes.

We did observe important benefits of co-design that fit this model. First and foremost was the shift to increased ownership and agency by the teachers. A classic failure in the diffusion of innovations occurs when change agents fail to grow advocacy from within the practitioner community. As noted above, participation in co-design transformed the teachers involved into advocates who became the ultimate spokespeople for the innovation within their school district. Because stakeholder advocacy is so crucial to adoption of new practices and new tools, this outcome of the co-design process should be considered when choosing among alternative design approaches.

Another potential benefit of co-design is teacher professional development. Our analysis shows that participation in co-design can promote teacher learning. The co-design process offered teachers a chance to develop and refine their own ideas about teaching in the framework of exploring how new software works in their classrooms. Two of our teachers reflected on multiple occasions about how they benefited from the opportunity to reflect on their teaching as part of the co-design process. One middle school teacher commented that researchers' perspectives were so different from teachers' that they gave her a new way to look at her own classroom teaching. She said that her design team meetings included "phenomenal conversations about what improves learning. The team is making me realize that I have always focused on classroom management, not on how the kids are actually learning. This affects my teaching tremendously."

An elementary teacher observed changes in how her students learned with the software, and these were changes she had not expected. When she introduced her students to software that collects students' questions, she noted:

> I was amazed at the impact of using questions as an assessment tool. When I had the kids ask questions, even on information we had covered and tested before, it really helped me to see the flaws in their understanding. We could then go back and discuss them. Also, there was a problem because they don't yet have good writing skills, so it was difficult for them to formulate questions.

9. Conclusion: Taking Co-Design Forward

Our study has attempted to define *co-design* and to map a possible terrain of the dynamics of co-design using participant experiences from a single project to illustrate those dynamics. In that sense, we believe our study may be instructive for future co-design efforts and research studies that aim to investigate systematically the impact of co-design on teachers and their practice. This study could help educational leaders make a more informed argument for co-design, establish more realistic expectations, and manage the emergent tensions among teachers and researchers as they work together toward an innovation goal.

At this point, one cannot make an argument for selecting co-design on the basis that it is more efficient or that it necessarily produces better-quality innovations or student learning outcomes. One can argue for selecting co-design because it powerfully surfaces and addresses the tensions between practitioners' and researchers' views of teaching and learning and thus has the potential to result in innovations that are both theoretically and practically compelling. Thus, co-design is one way to build community and common language among researchers and teachers and to bridge their views in concrete designs. Further, co-design can be a form of teacher professional development. Finally, co-design often results in teachers' taking increasing ownership and agency in the design and dissemination of the innovation, which makes broad-scale impact more likely.

Upon launching a co-design effort, the involved researchers and teachers need realistic mutual expectations of what will take place. Our study suggests it is particularly important for teachers to be aware of how slow and frustrating the early stages of design are likely to be, for example. On the basis of this study, it would be useful to launch a team by mapping the likely phases of a co-design process as they evolve over a school year. Our study also suggests that everyone involved should be realistic about the differentiation of roles that will be necessary to success; co-design does not proceed through pure democracy, although it can strive to maximize the voice of each participant in the emerging design. We would recommend that the leaders of a co-design review the tensions and process components we have highlighted, use these in their planning before a co-design process begins, and communicate their plan to the team.

Finally, awareness of evolving dynamics at different stages of co-design can help leaders and facilitators manage tensions towards successful resolution. We found that it is particularly important for leaders or facilitators to be aware of the time pressures on participants and also to work to give teachers equal voice. By being aware of how long it takes to arrive at a shared worldview among teachers and researchers, facilitators can avoid pushing too hard early on for design outcomes and honor the needs of all participants to build mutual trust and understanding. In the middle phase, when roles become more differentiated and the pace quickens, coordinating the work across parties becomes increasingly important. In the last phase, as agency and ownership begin to shift, facilitators could nurture emerging teacher leaders. These will be the people who carry the innovation forward into broader integration in their schools, districts, and larger educational systems.

Integration of innovative technologies into teaching and learning practices remains an elusive yet important goal at the federal, state, and local levels; integration is necessary for policymakers to demonstrate to constituents that investments in technology pay off — not only in terms of increasing access but also in terms of educational improvement. In our experience, the co-design process can play a role in creating a tighter integration of teaching practice, curriculum and technology. By defining *co-design*, describing key process steps, analyzing the tensions, and revealing a key dynamic of increasing ownership and agency, we have sought to contribute to refining and spreading this promising method of design.

Acknowledgments

An abbreviated version of this paper was presented at the 7th International Conference of the Learning Sciences. We would like to thank Barry Fishman and six anonymous reviewers for their comments on earlier versions of this manuscript. We would also like to acknowledge the members of the Project WHIRL co-design team for their help in making the project a success. This material is based in part on work supported by the National Science Foundation under Grant Number REC-0126197. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Atkin, J. M. (2001, April). *How science teachers improve their formative assessment practices.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Atkin, J. M., & Black, P. (2003). *Inside science education reform: A history of curricular and policy change.* New York: Teachers College Press.

Barab, S., & Luehmann, A. L. (2003). Building sustainable science curriculum: Acknowledging and accommodating local adaptation. *Science Education, 87*(4), 454–467.

Becker, H. J., & Anderson, R. J. (1998). *Teaching learning, and computing, 1998: A national survey of schools and teachers describing their best practices, teaching philosophies, and uses of technology.* Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Black, P., & Harrison, C. (2001). Feedback in questioning and marking: The science teacher's role in formative assessment. *School Science Review, 82*(301), 55–61.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74.

Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling up technology-embedded project-based science in urban schools. *Educational Psychologist, 35*(3), 149–164.

Bowers, J., & Pycock, J. (1994). Talking through design: Requirements and resistance in cooperative prototyping. In *Proceedings of the SIGCHI conference on human factors in computing systems: Celebrating interdependence* (pp. 299–305). Boston, MA: ACM.

Carr, A. A. (1997). User-design in the creation of human learning systems. *Educational Technology Research and Development, 45*(3), 5–22.

Carroll, J. M. (1995). *Scenario-based design.* New York: Wiley.

Cobb, P., Confrey, J., diSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

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

Cuban, L. (2001). *Oversold and underused: Computers in the classroom.* Cambridge, MA: Harvard University Press.

Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work.* New York: Teachers College Press.

Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences, 11*(1), 105–121.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *The Journal of the Learning Sciences, 8*(3&4), 391–450.

Ehn, P. (1992). Scandinavian design: On participation and skill. In P. S. Adler & T. A. Winograd (Eds.), *Usability: Turning technologies into tools* (pp. 96–132). New York: Oxford University Press.

Fishman, B. J., Penuel, W. R., & Yamaguchi, R. (2006). Fostering innovation implementation: Findings about supporting scale from GLOBE. In S. A. Barab, K. E. Hay & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences* (Vol. 1, pp. 168–174). Mahwah, NJ: Erlbaum.

Friedman, B. (Ed.). (1996). *Human values and the design of computer technology.* New York: Cambridge University Press.

Friedman, B., Kahn, P., & Borning, A. (2003). *Value sensitive design: Theory and methods.* Unpublished manuscript, Seattle, WA: University of Washington.

Garet, M. S., Porter, A. C., Desimone, L., 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.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Gorden, V. S., & Bieman, J. M. (1995). Rapid prototyping: Lessons learned. *Software, IEEE, 12*(1), 85–95.

Hand, V., Underwood, J., & Nielsen, L. (2001, April). *Teaming with teachers to integrate technology and curriculum: Lessons from the ESCOT Project.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Hughes, J., King, V., Rodden, T., & Andersen, H. (1995). The role of ethnography in interactive systems design. *Interactions*, 56–65.

Krippendorff, K. (2006). *The semantic turn: A new foundation for design.* New York: Taylor and Francis.

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.

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

Means, B., Penuel, W. R., & Padilla, C. (2001). *The connected school: Technology and learning in high school.* San Francisco, CA: Jossey-Bass.

Millen, D. R. (2000). Rapid ethnography: Time deepening strategies for HCI field research. In *Proceedings of the conference on designing interactive systems: Processes, practices, methods, and techniques* (pp. 280–286). New York, NY: ACM Press.

Muller, M. J., & Kuhn, S. (1993). Participatory design. *Communications of the ACM, 36*(6), 24–28.

Muller, M. J., Wildman, D. M., & White, E. (1993). Taxonomy of participatory design practices: A brief practitioner's guide. *Communications of the ACM, 36*(4), 26–28.

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

National Research Council (2001). *Classroom assessment and the National Science Education Standards.* Washington, DC: National Academy Press.

Penuel, W. R., Frank, K. A., & Krause, A. (2006). The distribution of resources and expertise and the implementation of schoolwide reform initiatives. In S. A. Barab, K. E. Hay & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences* (Vol. 1, pp. 522–528). Mahwah, NJ: Erlbaum.

Penuel, W. R., Tatar, D., & Roschelle, J. (2004). The role of research on contexts of teaching practice in informing the design of handheld learning technologies. *Journal of Educational Computing Research, 30*(4), 331–348.

Petrosino, A. J. (1998). *The use of reflection and revision in hands-on experimental activities by at risk children.* Unpublished dissertation. Vanderbilt University, Nashville, TN.

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.), *Fourth International Conference of the Learning Sciences.* (pp. 342–349). Mahwah, NJ: Erlbaum.

Roschelle, J., DiGiano, C., Koutlis, M., Repenning, A., Phillips, J., Jackiw, N. *et al.* (1999). Developing Educational Software Components. *IEEE Computer, 32*(9), 50–58.

Roschelle, J., Penuel, W. R., Yarnall, L., & Tatar, D. (2004, March). *Handheld tools that "informate" assessment of student learning in science: A requirements analysis.* Paper presented at the Second IEEE International Workshop on Mobile and Wireless Technologies in Education, JungLi, Taiwan.

Rosenholtz, S. (1989). *Teacher's workplace: The social organization of schools.* New York: Longman.

Rosenholtz, S., & Simpson, C. (1990). Workplace conditions and the rise and fall of teachers' commitment. *Sociology of Education, 63*, 241–257.

Sawyer, R. K. (2004). Creative teaching: Improvisation in the constructivist classroom. *Educational Researcher, 33*(2), 12–20.

Shepard, L. (1997). *Insights gained from a classroom-based assessment project.* CSE Technical Report 451. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Shumar, W. (2003). The role of community and belonging in online learning. In M. Mardis (Ed.), *Digital Libraries as Complement to K-12 Teaching and Learning.* Washington, DC: U.S. Department of Education.

Shrader, G., Williams, K., Lachance-Whitcomb, J., Finn, L.-E., & Gomez, L. (2001). *Participatory design of science curricula: The case for research for practice.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Snyder, C. (2003). *Paper prototyping: The fast and easy way to design and refine user interfaces.* San Francisco, CA: Morgan Kauffman.

Soloway, E., Guzdial, M., & Hay, K. E. (1994). Learner-centered design: The challenge for HCI in the 21st century. *Interactions, 1*(2), 36–47.

Stiggins, R. J. (1997). *Student-centred classroom assessment.* Columbus, OH: Merrill.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage.

Talbert, J. E., & McLaughlin, M. W. (2002). Assessing the school environment: Embedded contexts and bottom-up research strategies. In S. L. Friedman & T. D. Wachs (Eds.), *Measuring environment across the life span: Emerging methods and concepts.* (pp. 197–227). Washington, DC: APA.

Tripp, S. D., & Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. *Educational Technology, Research, and Development, 38*(1), 31–44.

White, B., & Frederiksen, J. R. (1998). Inquiry, modeling and meta-cognition: Making science accessible to all students. *Cognition and Instruction, 16*, 3–118.

Windschitl, M., & Sahl, K. (2002). Tracing teachers' use of technology in a laptop computer school: The interplay of teacher beliefs, social dynamics, and institutional culture. *American Educational Research Journal, 39*(1), 165–205.

Yarnall, L., Shechtman, N., & Penuel, W. R. (2006). Using handheld computers to support improved classroom assessment in science: Results from a field trial. *Journal of Science Education and Technology, 15*(1), 1–17.