20 Design Research

An Analysis and Critique

Paul Cobb

Vanderbilt University

Kara Jackson

University of Washington

Charlotte Dunlap

Vanderbilt University

In this chapter, we delineate the prototypical characteristics of the design research methodology and describe what is involved in conducting a design study to investigate either students' learning in a particular mathematical domain or teachers' development of increasingly sophisticated forms of practices. In addition, we discuss some of the common limitations of design studies, thereby identifying areas for attention in future studies of this type.

Design studies entail "engineering" participants' development of particular forms of practice while systematically studying the development of those practices and the context in which they emerge, which includes the designed means of support (Schoenfeld, 2006). Design studies are therefore both pragmatic and theoretical in orientation (Design-Based Research Collaborative, 2003). Pragmatically, they involve investigating and improving a design for supporting learning. Theoretically, they involve developing, testing, and revising conjectures about both learning processes and the means of supporting that learning (Gravemeijer, 1994b). The resulting theory then constitutes the rationale for the design.

Design studies can be conducted in a diverse range of settings that vary in type and scope. At one end of the spectrum, in *one-on-one design studies* a researcher conducts a series of individual teaching sessions with each of a small number of students in order to study the process of learning in a particular mathematical domain (e.g., Cobb & Steffe, 1983; Simon et al., 2010). At the other end of the spectrum, in *organizational design studies* a research team collaborates with teachers, school administrators, and other stakeholders to investigate and support the development of school and school district capacity for instructional improvement in mathematics (e.g., Cobb & Jackson, 2012; Fishman, Marx, Blumenfeld, & Krajcik, 2004). In this chapter, we focus on two common types of design studies:

- *Classroom design studies* in which a research team collaborates with a mathematics teacher (who might be a research team member) to assume responsibility for instruction in order to investigate the process of students' learning in a particular mathematical domain (e.g., Lamberg & Middleton, 2009; Lehrer & Kim, 2009; Simpson, Hoyles, & Noss, 2006; Stephan & Akyuz, 2012).
- *Professional development design studies* in which a research team works with a group of practicing mathematics teachers to support their development of increasingly sophisticated instructional practices (e.g., Cobb, Zhao, & Dean, 2009; Lesh & Kelly, 1997; Zawojewski, Chamberlin, Hjalmarson, & Lewis, 2008).

In general, it is appropriate to conduct these two types of design studies when research problems have the following two characteristics. First, the goal is to understand either how students develop specific forms of mathematical practice or how teachers develop particular forms of instructional practice. However, these developments rarely occur in situ and are therefore difficult if not impossible to study by conducting observational investigations. An interventionist methodology such as design research that aims to bring about the intended developments in order to study them is therefore appropriate. Second, current research on the process of supporting the development of the focal practice is inadequate and cannot inform the formulation of viable instructional or professional development designs. A bootstrapping methodology such as design research in which designs are improved in the course of iterative cycles of design and analysis is therefore appropriate.

Five crosscutting features characterize all types of design studies, including those that investigate students' and teachers' learning. Although some of these features are shared with other methodologies, when taken together they differentiate design research from other approaches. The first crosscutting feature is that, ideally, design studies address the types of problems that arise for practitioners as they attempt to support students' or teachers' learning, and thus contribute directly to improving the quality of educational practice.

The second feature is the highly interventionist nature of the methodology. The intent when conducting a design study is to investigate the possibilities for educational improvement by supporting either students' or teachers' development of relatively novel forms of practice in order to study their development. Consequently, the type of instruction or professional development enacted in the course of a study usually differs significantly from typical instructional or teacher education practice. The process of engineering the forms of learning being studied provides the research team with both considerable control compared with naturalistic investigations, and with the opportunity to identify forms of supports that are necessary for the development of the focal practices.

The third feature is that design studies have a strong theoretical as well as a pragmatic orientation. A primary purpose when conducting a design study is to develop theory that comprises substantiated conjectures about both processes of learning and the means of supporting that learning. These theories are modest in scope and focus on either students' development of particular types of mathematical reasoning in the classroom or teachers' development of particular forms of instructional practice in the context of professional development.

The fourth feature is that design studies involve testing and, if necessary, revising or abandoning conjectures about students' or teachers' learning processes and the means of supporting that learning. This process of testing and revising conjectures and thus of improving the associated design for supporting participants' learning involves iterative cycles of design and analysis. At any point in a design study, the evolving instructional or professional development design reflects then-current conjectures about the process of the participating students' or teachers' learning and the means of supporting it. Ongoing analyses of both the participants' activity and of the enacted supports for their learning provide opportunities to test, refine, and revise the underlying conjectures, and these revisions in turn inform the modification of the design.

The fifth crosscutting feature is that, as a consequence of the concern for theory, design studies aim for generalizability. Although a design study is conducted in a limited number of settings, the intent is not merely to investigate the process of supporting a particular group of students' or teachers' learning. Instead, the research team frames the initial design formulated when preparing for a study and the learning processes it is intended to support as an instance of a broader class of phenomena, thereby making them susceptible to theoretical analysis.

In the following sections, we first give a brief historical overview of the design research methodology to clarify its antecedents in both the learning sciences and mathematics education. We then discuss, in turn, classroom design studies and professional development design studies. For each type of study, we first consider a key research tool, the interpretive framework that the research team uses to make ongoing interpretations of participants' activity, and the enacted supports for their learning. For each, we then discuss the phases of preparing for a study, experimenting to support learning, and conducting retrospective analyses of data generated in the course of the study. Finally, we take a critical perspective by discussing some of the common limitations of each type of design study, thereby indicating areas for improvement in future studies of this type.

HISTORICAL OVERVIEW

The five defining features of design studies foreground the intimate relation between theory and practice. Methodologies in which instructional design serves as a context for the development of theories of learning and instruction have a long history, particularly in the former Soviet Union (Menchinskaya, 1969). However, the term ***design research*** emerged relatively recently in the learning sciences and is most closely associated with Ann Brown (1992) and Alan Collins (1992). In proposing design studies that investigated learning as it occurred in complex settings such as classrooms, Brown and Collins sought to overcome the perceived limitations of traditional studies of cognition that involve the control of variables in relatively artificial laboratory settings. They developed an analogy with design sciences such as aeronautical engineering to emphasize that the methodology is highly interventionist and has a theoretical as well as a pragmatic intent. As they explained, an aeronautical engineer creates a model that embodies theoretical conjectures, investigates how the model behaves under certain conditions, and generates data in order to test and revise the conjectures inherent in the model. Similarly, researchers conducting a design study create an initial design for supporting envisioned learning processes, investigate how the design plays out in practice, and generate data in order to test and revise theoretical conjectures inherent in the design.

Although design research in the learning sciences and in mathematics education is highly compatible, the histories differ. The emergence of the learning sciences from cognitive science signaled a relatively radical change of priorities (cf. DeCorte, Greer, & Verschaffel, 1996). In contrast, the development of design research in mathematics education has been more evolutionary and builds on two prior lines of research: the constructivist teaching experiment and Realistic Mathematics Education developed at the Freudenthal Institute in The Netherlands.

Steffe and his colleagues drew heavily on earlier Soviet work when they developed the constructivist teaching experiment methodology in the late 1970s and early 1980s (Cobb & Steffe, 1983; Steffe, 1983; Steffe & Kieren, 1994; Steffe & Thompson, 2000). The purpose of the teaching experiment as formulated by Steffe was to enable researchers to investigate the *process* by which individual students reorganize their mathematical ways of knowing. To this end, a researcher typically interacts with students one-on-one and attempts to precipitate their learning by posing theoretically motivated tasks and by asking follow-up questions, often with the intention of encouraging the student to reflect on her or his mathematical activity. The primary products of a constructivist teaching experiment typically consist of conceptual models composed of theoretical constructs that account for the learning of the participating students. Such constructs prove useful when accounting for the learning of other students, and can thus inform teachers' decision making (Thompson & Saldanha, 2000). Although the researcher acts as a teacher in this methodological approach, the primary emphasis is on the interpretation of students' mathematical reasoning rather than on the development of instructional designs.

Subsequent attempts to adapt the constructivist teaching experiment methodology to the classroom setting involved creating sets of instructional activities. However, the primary focus of these classroom experiments was on the development of explanatory constructs rather than the improvement of instructional designs (Cobb, Yackel, & Wood, 1995). For example, the intent of one series of analyses was to develop an interpretive framework that situated students' mathematical learning within the social context of the classroom (Cobb & Yackel, 1998). In retrospect, it is now apparent that a limitation in this work was the lack of specific, empirically grounded design heuristics that could inform the development of instructional activities.

The second line of research on which design research in mathematics education draws, Realistic Mathematics Education (RME), complemented the constructivist teaching experiment by focusing primarily on the design of instructional sequences rather than the development of explanatory theoretical constructs (cf. Gravemeijer, 1994b; Streefland, 1991; Treffers, 1987). RME researchers' work in developing, trying out, and modifying instructional sequences in a wide range of mathematical domains was oriented by Freudenthal's (1973) notion of mathematics as a human activity and informed by his didactical phenomenology of mathematics. The heuristics for instructional design in mathematics education that RME researchers proposed were delineated by reflecting on the process of designing and improving these specific instructional sequences (Gravemeijer, 1994a; Treffers, 1987).

It should be apparent from this brief historical account of design research that the initial focus was on supporting and investigating students' learning and that the methodology was only later extended to investigate teachers' learning. This historical overview also indicates that two types of conceptual tools are essential when conducting a design study to investigate and support either students' or teachers' learning: an interpretive framework for making sense of participants' activity in the complex settings in which design studies are conducted, and a set of design heuristics or principles that can guide the development of specific designs.

CLASSROOM DESIGN STUDIES

Interpretive Framework

In conducting a classroom design study, the research team makes ongoing interpretations of both the students' mathematical activity and the classroom learning environment. These interpretations necessarily involve suppositions and assumptions about mathematical learning processes and about the aspects of the classroom learning environment that are potentially important supports for students' learning. For example, some researchers assume that mathematical learning is a process of individual cognitive reorganization that occurs as students attempt to solve tasks and respond to the teacher's questions in the classroom (Clements & Sarama, 2004; Saldanha & Thompson, 2007). For these researchers, aspects of the classroom learning environment influence the process of students' learning by precipitating students' internal reorganization of their reasoning. Researchers who adopt this perspective on learning tend to foreground mathematical tasks together with physical, symbolic, and computer-based tools, and the teacher's questions, as key supports for students' learning. In contrast, other researchers assume that students' mathematical learning is situated with respect to classroom mathematical practices that are constituted collectively by the teacher and students (Doorman, Drijvers, Gravemeijer, Boon, & Reed, 2013; Kwon, Ju, Kim, Park, & Park, 2013; Lehrer, Kim, & Jones, 2011; Stephan & Akyuz, 2012). For these researchers, aspects of the classroom learning environment influence not merely the process of students' learning but its products, including the forms of mathematical reasoning that they develop. Researchers who take this latter perspective on learning typically focus on the affordances of classroom tasks and tools, and on the nature of classroom norms and the quality of classroom discourse as potential supports for students' learning.

A research team's suppositions and assumptions about mathematical learning are consequential because they influence ongoing design and instructional decisions. For example, Stephan and Akyuz (2012) conducted a classroom design study in which they supported the development of seventh-grade students' understanding of integers and the meaning of the minus sign. These concepts are typically included in elementary and middle-grades curricula but continue to prove problematic for students, in part because they are often reduced to procedures whose meaning can be lost (Byrnes, 1992). Stephan and Akyuz's (2012) design for supporting the participating students' learning involved problem scenarios that focused on monetary transactions and the use of the vertical number line as a tool for solving these problems. The authors drew on the emergent interpretive perspective that treats individual students' mathematical reasoning as acts of participating in collective classroom mathematical practices. Operationalizing this perspective involved strategic collection and analysis of classroom video data in addition to individual student interviews. Documenting shifts in classroom mathematical practice involved analyzing classroom discourse as both an indicator of individual mathematical reasoning and as a signal of what had become taken-as-shared within the class.

As a further example, Kwon et al. (2013) used Toulmin's (1958) scheme of argumentation as an overarching framework when they traced developments in how eighth-grade students justified their reasoning about geometric patterns. Using this framework, the researchers made interpretations of classroom events that informed the design of the various components of the classroom learning environment which were central to the design. Components included tasks that necessitated students' explicit justification of their reasoning, the organization of classroom activities, and the teacher's use of discursive moves that supported students' development of more complex and elaborated forms of argumentation. In addition, Kwon et al. used this framework when they conducted retrospective analyses of data generated during the study to connect the students' increasingly sophisticated mathematical arguments to the designed supports for their learning.

In our view, it is essential that researchers conducting a classroom design study make explicit the theoretical commitments inherent in their interpretive perspective, given the role of these commitments in orienting the design of supports for students' learning. By articulating the key constructs used when interpreting the students' mathematical activity and the classroom learning environment, the research team subjects these constructs to public debate and scrutiny. Classroom design studies conducted from a range of different perspectives can make important contributions. However, we also note that a considerable body of evidence has accumulated in the years since Brown's (1992) and Collins' (1992) pioneering work that indicates the forms of mathematical reasoning children and adults develop are shaped by the settings of their learning and, in particular, by the collective practices in which they participate while learning (Doorman et al., 2013; Hall, 2001; Hoyles, Noss, & Pozzi, 2001; Kwon et al., 2013).

Preparing for a Classroom Design Study

Specifying Goals for Students' Mathematical Learning

As we have indicated, classroom design studies are useful in testing and revising conjectures about students' development of domain-specific forms of reasoning that rarely occur in situ. In specifying the forms of mathematical reasoning that constitute the goals for students' learning, it is therefore critical to question how the mathematical domain under consideration is typically represented in curricula by identifying the central, organizing mathematical ideas. Clearly, any prior studies that have investigated the possibilities for students' mathematical learning in the focal mathematical domain are relevant in this regard. A significant number of classroom design studies have been conducted that focus on elementary domains such as early number, whereas the relevant research base of some secondary and university level domains is extremely thin.

Although the formulation of student learning goals might also be informed by national or state policy documents that detail standards for students' mathematical learning, the goals proposed for a design study typically involve a significant reconceptualization of the relevant standards. For example, the learning goals that Stephan and Akyuz (2012) formulated for their seventh-grade design study in which they investigated students' learning in the domain of integers departed from then-current state standards by emphasizing how students come to reason quantitatively about integers.

In addition to taking account of policy recommendations, the delineation of learning goals might draw on analyses of the disciplinary practice of professionals. For example, Lehrer, Schauble, Strom, and Pligge (2001) discussed a series of design studies conducted in elementary classrooms in which they first supported students' development of mathematical models that involved similarity and ratio so that students could then investigate the volume, weight, and density of different types of materials. They explained that their decision to introduce mathematical modeling and the investigation of physical attributes of materials sequentially rather than simultaneously was in response to "typical forms of integration" (p. 43) of mathematical tools and scientific contexts, which often underestimate the power of students' conceptual reasoning about the former and reduce the latter to an inauthentic context for employing procedures. As a consequence, the design studies they conducted aimed to support students' development of practices that are nearly invisible in school science but are central to the work of practicing scientists: using mathematics as a tool to reason about and model differences in the properties of materials. As a further example, Cheeseman, McDonough, and Ferguson (2012) challenged the contention that the flexible use of formal units for measuring mass is beyond first-grade students. Their design study resulted in an instructional sequence that took students from concrete and comparative heuristic measurement routines to more formalized use of tools and units, and resulted in student learning about mass and measurement typically reserved for students in older grade levels. In both of these examples, the learning goals that oriented the entire instructional design effort took account of disciplinary practice.

Documenting Instructional Starting Points

In addition to specifying explicit learning goals, it is also important to identify the aspects of students' current reasoning on which instruction can build before attempting to formulate conjectures about students' development and the means of supporting it. Prior research, such as interview and observational studies, can be useful in indicating students' initial reasoning. However, it is often necessary to create additional forms of assessments when preparing for a design study, especially if little prior work has been conducted in the relevant domain or if the proposed learning goals differ significantly from those addressed by typical instruction. These assessments usually take the form of one-on-one interviews but might also involve observations of students as they attempt to reason through tasks. In addition, written assessments can be used if the research base is strong enough to guide the development of tasks that are aligned with the overall intent of the study. For example, Stephan and Akyuz (2012) used interviews and written assessments both to identify instructional starting points and to complement classroom observations as a way of tracking developments in the participating students' reasoning during the study.

Delineating an Envisioned Learning Trajectory

The next step in preparing for a classroom design study is to develop an initial design by specifying an envisioned or hypothetical learning trajectory that comprises testable conjectures about both significant developments in students' reasoning and the specific means of supporting these developments (Simon, 1995). In this regard, it is important to clarify the design heuristics or principles that informed the development of the initial design. For example, Stephan and Akyuz's (2012) envisioned trajectory for students' learning was informed by RME and included explicit conjectures about how students' reasoning about integers might develop as they used a series of increasingly sophisticated symbolic tools to solve a sequence of instructional tasks. Similarly, Wawro, Rasmussen, Zandieh, and Larson (2013) drew on RME as they formulated their initial design for supporting college students' learning in linear algebra. For their part, Lehrer et al.'s (2001) initial conjectures included possible benchmarks in students' developing reasoning about volume and similarity together with possible types of tasks and symbolic tools. Their design of instructional tasks was informed by the heuristic that mathematical models of scientific phenomena are analogies that show and hide (or distort) different aspects of a reality (cf. Hesse, 1965). The instructional sequence that they tested and improved exploits this "mismatch, or residual between the model and the world" (Lehrer et al., 2001, p. 52) in order to support students in formulating, critiquing, and revising models of the relations between volume, weight, and density.

It is worth noting that the intent when assessing the potential of particular types of tasks and of physical or symbolic tools is to anticipate the student learning opportunities that might arise if they were to be used in the classroom. In our view, it is therefore essential to envision how the tasks and tools might actually be enacted in the classroom by considering the nature of classroom norms and discourse (Gravemeijer & Cobb, 2006). This attention to the means of support sets an envisioned learning trajectory apart from the notion of a developmental trajectory as typically used in cognitive and developmental psychology by underscoring that the envisioned developments will not occur unless appropriate supports are enacted in the classroom. It is in this sense that the forms of learning being investigated are "engineered" in the course of a design study.

In our experience, prior studies that are useful in informing the delineation of an envisioned learning trajectory focus on learning goals that are at least partially compatible with those of the planned study and include reports of the process of students' learning, the instructional setting, and the supports for that learning. Because the number of such studies is limited in many domains, the initial conjectures about students' learning and the means of supporting it are often provisional and eminently revisable. The process of formulating the envisioned learning trajectory is nonetheless valuable because the research team is then in a position to improve its initial design in a data-driven manner once it begins experimenting in the classroom.

Placing the Study in Theoretical Context

An overriding goal when conducting a classroom design study is to produce "humble theory" (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003) that can provide others with useful guidance as they attempt to support students' learning in other settings. It is therefore critical when preparing for a design study to place it in a larger theoretical context by framing it as a paradigmatic case of a broader class of phenomena. For example, Stephan and Akyuz (2012) and Lehrer et al. (2001) both sought to develop a ***domain-specific instructional theory***. Stephan and Akyuz (2012) framed their study as a case of supporting the development of middle-grades students' reasoning about integers, and Lehrer et al. (2001) framed their study as a case of supporting elementary students' learning in particular mathematical and science domains.

These illustrations do not, of course, exhaust the possibilities. A series of design studies can, for example, serve as the context in which to revise and refine an initial interpretive framework that does useful work in informing the generation, selection, and assessment of design alternatives. Examples of frameworks developed in this way include the theory of meta-representational competence (diSessa, 1992, 2002, 2004), the theory of quantitative reasoning (Smith & Thompson, 2007; Thompson, 1994, 1996), the theory of actor-oriented abstraction (Lobato, 2003, 2012), and the emergent perspective on students' mathematical learning in the social context of the classroom (Cobb, Stephan, McClain, & Gravemeijer, 2001). In each of these cases, the framework was revised in response to issues encountered while using it to make sense of classroom events. As a consequence, the resulting framework does not stand apart from the practice of experimenting to support learning but is instead grounded in it. Such frameworks can function both as a source of guidance for instructional design and as tools for making sense of what is happening in the complex setting of the classroom (diSessa & Cobb, 2004).

Experimenting to Support Learning

The objective when conducting any type of design study is not to demonstrate that the envisioned learning trajectory works. The primary goal is not even to assess whether it works, although the research team will necessarily do so. Instead, the purpose when experimenting to support learning is to improve the envisioned trajectory developed during the preparation phase of the study. Improvement happens through testing and revising conjectures about both students' prospective learning processes and the specific means of supporting it.

Data Collection

Decisions about the types of data that need to be generated in the course of a study depend on the theoretical intent of the design study. The data have to make it possible for the research team to address the broader theoretical issues of which the learning setting under investigation is a paradigm case when subsequently conducting retrospective analyses. At a minimum, the research team has to collect data that allows them to document both the process of students' learning in the classroom sessions and the evolving classroom learning environment, which includes the enacted supports for the students' learning. Thus, as we have noted, Stephan and Akyuz (2001) conducted pre- and post-interviews and also used written assessments to assess shifts in individual students' reasoning about integers. Additionally, they video-recorded all classroom sessions and made copies of all the students' written work so that they could document the classroom mathematical practices that were established in the course of the study. The analysis of these data allowed the research team to investigate how the students' participation in successive practices both supported and constrained the development of the students' reasoning about integers.

Existing instruments are often not adequate because classroom design studies typically aim at novel learning goals. As a consequence, the research team usually has to devise ways of documenting the students' developing reasoning and key aspects of the classroom learning environment. The data collected in the course of a classroom design study are therefore usually qualitative for the most part. For example, one of the goals of Lehrer et al.'s (2001) study was to investigate how students might come to reason about density quantitatively. It would have been relatively straightforward to develop a pencil-and-paper assessment of students' proficiency in executing procedures for finding and comparing the densities of different materials. However, the challenge of documenting how the students' reasoning about density developed in relation to the designed classroom learning environment required that the research team conduct video-recorded one-on-one pre- and post-interviews with the participating students, video-record all classroom sessions, and make copies of all the artifacts that the students produced during the study.

Iterative Cycles of Design and Analysis

The iterative nature of a design study is a key aspect of the methodology. Each cycle involves designing instruction, enacting that design during a classroom session, and then analyzing what transpired in the classroom in order to plan for upcoming sessions. The overall goal in enacting successive design and analysis cycles is to test and improve the envisioned learning trajectory formulated during the preparation phase. As part of this testing and revision process, it is essential to have debriefing meetings after each classroom session in which members of the research team share and debate their interpretations of classroom events. Once the team has reached consensus, it can then prepare for upcoming classroom sessions by designing (or revising existing designs for) instructional tasks and considering other means of support (e.g., the renegotiation of classroom norms).

It is also useful to have longer research team meetings periodically in order to take stock of the ongoing process of testing and revising conjectures. The purpose of these meetings is to outline a revised learning trajectory for the entire study that takes account of the revisions made thus far. In our view, ensuring that there is a reflexive relationship between local judgments (e.g., the specific tasks that will be used in a particular session and the mathematical issues on which the teacher might press students) and the longer term learning goals and overall learning trajectory should be a basic tenet of design research (Simon, 1995).

Conducting Retrospective Analysis

The final phase of a design study involves conducting retrospective analyses by drawing on the entire data set generated while experimenting in the classroom. The ongoing analyses conducted while the study is in process usually relate directly to the immediate pragmatic goal of supporting the participating students' learning. In contrast, retrospective analyses seek to place this learning and the means by which it was supported in a broader theoretical context by framing it as a paradigmatic case of a more encompassing phenomenon. For ease of explication, we assume that one of the primary theoretical goals of a classroom design study is to develop a domain-specific instructional theory.

Kelly (2004) observed that methodologies are underpinned by distinct argumentative grammars that link research questions to data, data to analysis, and analysis to final claims and assertions. He noted that the argumentative grammar of mature methodologies, such as randomized field trials, can be described separately from the details of any particular study, and then went on to observe that there is no agreed-upon argumentative grammar for design research. As a consequence, "design studies lack a basis for warrant for their claims" (p. 119). This is clearly a severe weakness of the methodology. We therefore propose an argumentative grammar for classroom design studies and then discuss issues of trustworthiness specific to classroom design studies in the following paragraphs.

Argumentative Grammar

The first step in the proposed argumentative grammar is to demonstrate that the students would not have developed the documented forms of mathematical reasoning but for their participation in the design study. Assuming that sound procedures have been employed to assess developments in the students' reasoning, this step in the argument is usually straightforward because classroom design studies aim to investigate students' development of novel forms of reasoning that rarely emerge in the context of typical mathematics instruction. The team can therefore draw on prior interview and observational studies to show that the documented forms of reasoning are relatively rare. As Brown (1992) made clear, the suggestion that the students' learning can be attributed to the Hawthorne Effect is not viable because the research team has predicted the forms of reasoning the students would develop when preparing for the study.

The second, more demanding, step in the proposed argumentative grammar is to show that the findings are potentially generalizable by delineating the aspects of the investigated learning process that can be repeated in other settings. This concern for replicability does not imply that a design should be realized in precisely the same way in different classrooms. Instead, the intent is to inform others as they customize the design to the settings in which they are working by differentiating between the necessary and the contingent aspects of the design. A primary concern when conducting a retrospective analysis of the entire data corpus is to document how each successive form of reasoning emerged as a reorganization of prior forms of reasoning, and to identify the aspects of the classroom learning environment that supported the students' development of these successive forms of reasoning. The resulting domain-specific instructional theory explains how the students' learning was engineered by explicating what Brown (1992) characterized as the coupling between successive developments in their reasoning and the relevant aspects of the classroom learning environment, including the designed supports as they were enacted in the classroom. The likelihood that the research team will be able to construct a robust theory of this type is greater if it takes a broad view of possible supports that extends beyond instructional tasks and tools, and if it employs an interpretive framework that treats students' mathematical learning as situated with respect to the classroom learning environment.

It is important to note that in explaining how students' learning was supported in the design study classroom, retrospective analyses of the type that we have outlined differentiate the necessary aspects of the classroom learning environment from those that are contingent and might be varied by researchers working in other settings. For example, the sequence of instructional tasks that Stephan and Akyuz (2012) used took the notion of net worth as a grounding context for reasoning about positive and negative integers. Their retrospective analysis indicated that this task context directly supported and was necessary for the students to come to reason about integers quantitatively. Their analysis also indicated that the students' use of particular symbolic tools (such as the vertical number line), teacher press on particular issues (e.g., "Who is worth more?"), and students' use of certain gestures to indicate differences and changes in quantities were also necessary, whereas the specific number combinations used in tasks were contingent and might be varied by others building on their work.

It should be clear that the generalizability of the findings of a design study is not based on a representative sample and what Maxwell (2004) called a *regularity type of causal description* that captures observed regularities across a number of cases. Instead, it is based on a *process-oriented explanation* of a single case "that sees causality as fundamentally referring to the actual causal mechanisms and processes that are involved in particular events and situations" (Maxwell, 2004, p. 4). In this regard, Maxwell drew on Shadish, Cook, and Campbell (2002) to clarify that process-oriented explanations are concerned with "the mechanisms through which and the conditions under which the causal relationship holds" (p. 4). In the case of a domain-specific instructional theory, the mechanisms are the processes by which specific aspects of the learning environment support particular developments in students' reasoning, and the conditions are the students' reasoning at a particular point in a learning trajectory.

In summary, the argumentative grammar that we have outlined involves:

- Demonstrating that the students would not have developed particular forms of mathematical reasoning but for their participation in the design study.
- Documenting how each successive form of reasoning emerged as a reorganization of prior forms of reasoning.
- Identifying the specific aspects of the classroom learning environment that were necessary rather than contingent in supporting the emergence of these successive forms of reasoning.

In presenting this argumentative grammar, we have spoken as though a robust instructional theory can be developed in the course of a single study. However, this is not always the case, especially if the research base on which the team can build when formulating initial design conjectures is thin. Instead, it is sometimes necessary to conduct a series of studies in which the findings of one study inform the initial design for the next study (Gravemeijer & Cobb, 2006). For example, the domain-specific instructional theories that Stephan and Akyuz (2012) and Lehrer et al. (2001) developed were refined while conducting a series of studies. Even when a single study does appear to be sufficient, we believe it is useful to conduct follow-up trials with a range of participants in a variety of settings. These trials are not necessarily full-scale design studies but focus on customizing the design while working in a new setting.

Trustworthiness

Trustworthiness is concerned with the reasonableness and justifiability of claims and assertions about both successive developments in the participating students' reasoning and the aspects of the classroom learning environment that supported those developments. Clearly, a discussion of the basic tenets of qualitative data analysis is beyond the scope of this chapter. However, we should acknowledge that analyzing the large longitudinal data set generated in the course of a classroom design study can be challenging. It is nonetheless essential to analyze the entire data corpus systematically while simultaneously documenting all phases of the analysis process, including the evidence for particular inferences. Only then can final claims and assertions be justified by backtracking through the various levels of the analysis, if necessary, to the original data sources (e.g., video recordings of classroom sessions and audio-recorded student interviews). It is the documentation of the research team's data analysis process that provides an empirical grounding for the analysis. The documentation of this process enables other researchers to differentiate systematic analyses in which sample episodes are used to illustrate general assertions from untrustworthy analyses in which a few possibly atypical episodes are used to support unsubstantiated claims. Additional criteria that enhance the trustworthiness of a retrospective analysis include both the extent to which it has been critiqued by other researchers who do not have a stake in the success of the study, and the extent to which it derives from a prolonged engagement with students and teachers (Taylor & Bogdan, 1984). This latter criterion is typically satisfied in the case of classroom design studies and constitutes a strength of the methodology.

PROFESSIONAL DEVELOPMENT DESIGN STUDIES

As we have indicated, classroom design studies are frequently conducted to develop domain-specific instructional theories that consist of:

- A substantiated learning process that culminates with students' attainment of significant learning goals in a particular mathematical domain.
- The demonstrated means of supporting that learning process.

Similarly, a primary goal when conducting a professional development design study is to develop what we call a *practice-specific professional development theory* that consists of:

- A substantiated learning process that culminates with mathematics teachers' development of particular forms of instructional practice.
- The demonstrated means of supporting that learning process.

Pragmatically, professional development design studies involve supporting teachers in improving specific aspects of their instructional practice. Following Ball and Cohen (1999), we take it as given that teacher professional development should center on "the critical activities of the profession" and "emphasize question, investigations, analysis, and criticism" (p. 13). Theoretically, professional development design studies involve developing, testing, and revising conjectures about both the process by which teachers develop increasingly sophisticated instructional practices and the means of supporting that development. In this regard, Grossman, Compton, et al. (2009) observed that "practice in complex domains involves the orchestration of skill, relationship, and identity to accomplish particular activities with others in specific environments" (p. 2059). As a consequence, the conjectures about teachers' learning are not restricted to directly observable aspects of teaching (e.g., questioning students) but can include a focus on the development of particular types of knowledge (e.g., knowledge of students' mathematical reasoning in a particular domain) and beliefs (e.g., beliefs about the mathematical capabilities of particular groups of students) that are implicated in the enactment of particular instructional practices (cf. Bannan-Ritland, 2008).

As a point of clarification, we use the term *professional development* (PD) to refer to activities that are intentionally designed to support teachers' learning. PD therefore includes both pull-out sessions for teachers from a number of schools or from a single school that are led by a researcher or by a facilitator who is a member of the research team, and one-on-one support in which a researcher or a coach who is a member of the research team works with individual teachers in their classrooms.

The number of PD design studies that have been conducted is relatively small compared with classroom design studies. However, many of the basic tenets of PD design studies parallel those of classroom design studies. We will therefore focus primarily on the instances in which tenets have to be modified significantly and on additional issues that need to be addressed when conducting a PD design study.

Interpretive Framework

The interpretive framework that a research team uses when conducting a PD design study explicates its suppositions and assumptions about the process of teachers' learning and about aspects of the PD learning environment that are necessary rather than contingent in supporting that learning. This framework should address two issues that do not typically arise when conducting classroom design studies: situating participants' activity with respect to school settings and accounting for the relations between their activity across two settings.

Situating Teachers' Activity with Respect to School Settings

A key difference between classroom design studies and PD design studies concerns the extent to which it is possible (and desirable) to insulate participants from the requirements and expectations of their schools. In classroom design studies, researchers typically isolate the study classroom to the greatest extent possible when negotiating entrée to the site. In contrast, it is usually not possible to renegotiate the school settings in which the participating teachers work when developing sites for a PD design study. We view this as an advantage given that the influence of professional development on what teachers do in their classrooms is mediated by the school settings in which they teach (e.g., Bryk, Sebring, Allensworth, Luppesco, & Easton, 2010; Cobb, McClain, Lamberg, & Dean, 2003; Coburn, 2003; Grossman, O'Keefe, Kantor, & Delgado, 2013). Key aspects of school settings include the instructional materials and associated resources to which teachers have access and that they are expected to use (e.g., pacing guides and curriculum frameworks), the people to whom teachers are accountable and for what they are held accountable (e.g., school principals' expectations for mathematics instruction), and the formal and informal sources of support on which teachers can draw to improve their instructional practices (e.g., school and district PD, colleagues to whom they can turn for advice about instruction).

In light of this difference between classroom and PD design studies, it is important that the interpretive framework a research team uses when conducting a PD study situates the participating teachers' activity with respect to key aspects of the settings in which they work (cf. Zawojewski et al., 2008). For example, the framework that Cobb, McClain, et al. (2003) used in a PD design study that focused on teaching statistical data analysis in the middles grades drew primarily on Wenger's (1998) theoretical analysis of communities of practice. This approach involved documenting the practices of members of distinct communities that had a stake in middle-grades mathematics teaching and learning (e.g., school leaders, district math leaders), and analyzing the connections between communities in terms of boundary encounters, brokers, and boundary objects. This attention to the school settings in which the teachers worked resulted in greater explanatory power when accounting for the teachers' activity in both PD sessions and their classrooms. This in turn enabled the research team to adjust their PD design accordingly.

Situating Teachers' Activity with Respect to the PD Learning Environment

It is also important that the interpretive framework the research team uses situates teachers' activity with respect to the PD learning environment, including the social norms established in the sessions, the PD activities in which they engage, the tools they use, and the terminology and discourse constituted during sessions. The research team's design efforts should focus on these aspects of the PD learning environment, as a considerable body of evidence indicates that they influence the practices and associated forms of reasoning that the participating teachers develop (e.g., Horn, 2005; Kazemi & Franke, 2004; Putnam & Borko, 2000; Sherin & Han, 2004). The research teams' assumptions about teachers' learning and the PD learning environment are therefore consequential because they influence ongoing design decisions.

Accounting for the Relations Between Teachers' Activity Across Settings

In a classroom design study, the research team typically focuses on supporting students' learning within a single setting, the classroom. In contrast, the intent of a PD design study is to engage teachers in activities in one setting, the PD sessions, with the explicit goal of supporting their reorganization of their activity in another setting, the classroom. As a consequence, designs for supporting teachers' learning necessarily involve suppositions and assumptions about the relations between teachers' activity across these two settings (Cobb, Zhao, et al., 2009; Kazemi & Hubbard, 2008). As Kazemi and Hubbard (2008) and Cobb, Zhao, et al. (2009) observe, PD has traditionally reflected an assumption that the relation between a PD session and a teacher's classroom is unidirectional; teachers' activity in PD sessions is assumed to impact what they do in their classrooms. However, a model of practice-based PD challenges this assumption by proposing that teachers' ongoing practice serves as an important resource for teachers' learning (Cobb, Zhao, et al., 2009; Kazemi & Hubbard, 2008).

The assumptions a research team makes about the relations between teachers' activity across the two settings impacts both the design and the interpretation of teachers' activity. It is therefore important for researchers conducting PD design studies to be explicit about how they conceptualize these relations. For example, in the PD design experiment that focused on statistical data analysis, Cobb, Zhao, et al. (2009) found that although the participating teachers readily analyzed student work in PD sessions, they did not view this activity as relevant to their classroom practice. It subsequently became apparent that while the research team assumed the teachers would view student work as a "resource for the prospective planning of future instruction" (p. 188), they used it in the classroom solely for "retrospective assessment." This finding led the research team to modify their interpretive framework to take account of how artifacts were used in each setting, and how each influenced the other. They were then in a position to adjust their design for supporting the teachers' learning to take account of this relation.

Preparing for a PD Design Study

To avoid repetition, we take our discussion of classroom design studies as a point of reference and limit our discussion to issues that are specific to PD design studies.

Specifying Goals for Teachers' Learning

Parallel to classroom design studies, PD design studies are useful in testing and revising conjectures about teachers' development of forms of instructional practice that rarely occur in situ and for which viable designs do not currently exist for supporting the development of the focal practices and associated forms of beliefs and forms of knowledge. The forms of instructional practice that constitute the goals of a PD design study should be specified in as much detail as possible to orient the formulation of an initial design. In our view, it is essential that the targeted forms of practice can be justified in terms of student learning opportunities. This implies that the first step in delineating the goals for teachers' learning is to clarify goals for students' mathematical learning (e.g., develop conceptual understanding as well as procedural fluency, explain and justify solutions, make connections among multiple representations). The second step is to then draw on current research on mathematics teaching to identify instructional practices that have been shown to support students' attainment of these mathematics learning goals. For example, current research suggests that if students are to develop conceptual understanding as well as procedural fluency, it is important that teachers routinely pose and maintain the rigor of cognitively demanding tasks (Henningsen & Stein, 1997), elicit and build on student thinking to advance an instructional agenda (Franke, Kazemi, & Battey, 2007), and orchestrate whole-class discussions in which students are pressed to make sense of each other's solution in relation to important mathematical ideas (Stein, Engle, Smith, & Hughes, 2008). These (and other) findings can inform the specification of goals for teachers' learning.

Documenting Instructional Starting Points

In addition to specifying explicit learning goals, it is important to identify aspects of teachers' current practices and relevant forms of knowledge on which PD can build before attempting to formulate conjectures about teachers' development of the target forms of practice and the means of supporting that development. Determining which instructional starting points to document will depend on the goals of the study. For example, it may be important to document the participating teachers' mathematical knowledge for teaching (Hill, Schilling, & Ball, 2004) or their conceptions of the mathematical capabilities of traditionally underserved groups of students (Jackson & Gibbons, 2014). Documenting starting points usually involves classroom observations, assessments (of mathematical knowledge for teaching, for example), and interviews. In addition, it is important (though atypical) to document the school settings in which the participating teachers work as these settings will mediate the influence of the PD on the participating teachers' classroom practices.

Delineating an Envisioned Learning Trajectory

The next step in preparing for an experiment is to delineate an envisioned learning trajectory by formulating testable conjectures about significant developments in teachers' classroom practices, knowledge, and beliefs, and the means of supporting these developments. In doing so, it is necessary to consider how their learning in the PD sessions might relate to changes in their classroom practices as they are situated in the school settings in which they work. The current literature on teacher learning and on professional development includes only a few analyses that report actual trajectories of mathematics teachers' development of particular forms of practice (see, e.g., Franke, Carpenter, Levi, & Fennema, 2001; Kazemi & Franke, 2004; van Es & Sherin, 2008). However, the literature on teacher learning and on professional development is useful in suggesting potentially productive means of supporting teachers' learning. For example, there is some evidence that in-service teacher PD that impacts classroom instruction shares the following qualities: it is sustained over time, involves the same group of teachers working together, is focused on issues central to instruction, and is organized around the instructional materials that teachers use in their classrooms (Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009; Garet, Porter, Desimone, Birman, & Yoon, 2001; Kazemi & Franke, 2004; Little, 2003).

In our view, the findings of recent research on practice-based preservice teacher education (e.g., Ball, Sleep, Boerst, & Bass, 2009; Lampert, Beasley, Ghousseini, Kazemi, & Franke, 2010; Lampert et al., 2013; McDonald, Kazemi, & Kavanagh, 2013) is particularly relevant in informing the design and enactment of supports for in-service teachers' learning and merits further investigation in the context of in-service teacher PD. This body of research, which is grounded theoretically in analyses of how professionals develop complex forms of practice, suggests it is crucial that teachers are provided opportunities to engage in both pedagogies of investigation and enactment (Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009) that are organized around target instructional practices (e.g., eliciting and building on student thinking to accomplish an instructional agenda). Pedagogies of investigation involve analyzing and critiquing representations of practice such as student work and video-cases of teaching (Borko, Jacobs, Eiteljorg, & Pittman, 2009; Kazemi & Franke, 2004; Sherin & Han, 2004). Pedagogies of enactment involve planning for, rehearsing, and enacting aspects of practice in a graduated sequence of increasingly complex settings (e.g., teaching other teachers who play the role of students, working with a small group of students, teaching an entire class). Opportunity for teachers to co-participate in activities that approximate the targeted practices with more accomplished others are crucial to pedagogies of enactment (Bruner, 1996; Forman, 2003; Lave & Wenger, 1991).

In addition to indicating potentially productive types of PD activities, the current literature on teacher learning and professional development can inform the practices of facilitators in leading the enactment of those activities that can be justified in terms of teacher learning opportunities. In our view, the recent developments in preservice teacher education that we have cited are particularly promising for in-service teacher education because they reflect a bidirectional view of the relation between teachers' activity in PD sessions and their classrooms. However, because research on the processes by which teachers develop particular forms of practice is relatively thin, initial design conjectures will almost certainly be provisional and thus eminently revisable.

Placing the Study in Theoretical Context

The intent of a PD design study is to produce knowledge that will be useful in providing guidance to others as they attempt to support teachers' learning in other settings. As is the case with classroom design studies, it is therefore important to frame a PD design study explicitly as a paradigmatic case of a broader class of phenomena, for example, teachers' development of particular practices (e.g., eliciting and responding to student thinking), knowledge (e.g., of students' reasoning in a particular mathematical domain), and/or beliefs (e.g., about the mathematical capabilities of particular groups of students).

Experimenting to Support Learning

The objective when experimenting to support the participating teachers' learning is to improve the envisioned trajectory by testing and revising conjectures about both the prospective learning processes and the specific means of supporting it.

Data Collection

The data collected have to make it possible for the research team to address the broader theoretical issues under investigation when conducting retrospective analyses. At a minimum, researchers will need to collect data to document:

- Relevant aspects of the school context that might mediate the impact of the PD on teachers' resulting practices, knowledge, and/or conceptions.
- Relevant aspects of the PD learning environment, including the enacted supports for the participating teachers' learning.
- The process of the teachers' learning in the PD sessions.
- Developments in the teachers' classroom practices.

The data are likely to be primarily qualitative, given the focus on accounting for the *process* by which teachers learn, and the means by which it is supported. However, it is reasonable to include validated quantitative instruments, such as an assessment of teachers' mathematical knowledge for teaching (Hill, Ball, & Schilling, 2008), if they fit with the theoretical intent of the study and contribute to the research team's understanding of teachers' ongoing learning (cf. Lesh & Kelly, 1997).

Cycles of Design and Analysis

In PD design studies, the cycles comprise a PD session together with the researchers' debriefing meetings held after the meeting to conduct an initial (ongoing) analysis of what transpired and to plan for future sessions. The cycles depend on the frequency of the PD sessions and are therefore usually less frequent than the daily cycles of a classroom design study. In the debriefing meetings, it is important to take account of the school settings in which the teachers work when developing explanations of their activity in PD sessions as well as of their classroom instruction.

Conducting Retrospective Analysis

As is the case for classroom design studies, ongoing analyses conducted while a PD design study is in progress contribute to the immediate pragmatic goal of supporting the participating teachers' learning, whereas retrospective analyses treat the teachers' learning and the means by which it was supported as paradigmatic of a more encompassing phenomenon. In discussing retrospective analyses, we assume that one of the primary goals of the study is to develop a practice-specific PD theory.

Argumentative Grammar

The argumentative grammar that we propose for PD design studies parallels that for classroom design studies. The two major differences concern accounting for changes in the teachers' activity across settings, and the need to take account of the mediating role of the school context. Although trustworthiness is central to the argumentative grammar for PD design studies, we do not discuss it explicitly because the issues addressed when discussing classroom design studies apply equally to PD design studies.

Similar to classroom design studies, the first step in the proposed argumentative grammar is to demonstrate that the teachers would not have developed the documented forms of instructional practice but for their participation in the design study. This is usually straightforward because PD design studies typically investigate teachers' development of forms of instructional practice that rarely occur in situ and for which viable designs do not currently exist.

The second step in the argumentative grammar involves showing that the findings are potentially generalizable. This can be accomplished by delineating the aspects of the investigated learning processes and the means of supporting them that are necessary rather than merely contingent, and by reporting how the school settings in which the teachers worked mediated the influence of the PD on their classroom practices. Similar to classroom design studies, this does not imply that a design should be repeated with absolute fidelity. Instead, the intent is to inform others of the necessary aspects of the PD design and of the school settings so that they can customize the design to the settings in which they are working. This entails conducting an analysis of the entire data corpus to document how the teachers developed increasingly sophisticated forms of instructional practice, to identify aspects of the PD learning environment that supported the teachers' development of these practices, and to clarify the mediating role of the school settings in which the teachers worked.

The resulting practice-based professional development theory explains how the teachers' learning was engineered by specifying relations between successive developments in teachers' practice and the relevant aspects of the PD learning environment and the school settings. The likelihood that the research team will be able to construct a robust theory of this type is greater if they take a broad view of the school settings in which teachers work and if they employ an interpretive framework that treats teachers' learning as situated with respect to the PD learning environment and school settings.

Similar to classroom design studies, the generalizability of the findings of a PD design study is based on a process-oriented explanation, in which the mechanisms through which and the conditions under which teachers developed the documented forms of practice are reported. In the case of a practice-specific professional development theory, the mechanisms are the process by which specific aspects of the PD learning environment support teachers' successive reorganizations of their practices. The conditions are the teachers' practices at a particular point in the substantiated learning trajectory and specific aspects of the school settings.

In summary, the argumentative grammar for a PD design study involves:

- Demonstrating that the participants would not have developed particular forms of practice but for their participation in the design study.
- Identifying the specific aspects of the PD learning environment that were necessary rather than contingent in supporting the emergence of these successive forms of practice.
- Clarifying how specific aspects of the school settings mediated the influence of the teachers' learning in PD sessions on their classroom practice.

Given that the research base on which a team can build when formulating initial design conjectures for a PD design study is thin, it is unlikely that a robust practice-specific PD theory can be developed in the course of a single study. Therefore, as is the case for classroom design studies, it is probably necessary to conduct a series of studies in which the findings of one study inform the initial PD design for the next study.

As an additional observation, much of what we have said about investigating and supporting teachers' learning can be generalized to the investigation of the learning of members of other role groups whose practices are implicated in school and district instructional improvement efforts, including mathematics coaches, school leaders, or professional development facilitators. The goal of such studies would be to develop a practice-specific professional development theory that consists of:

- A substantiated learning process that culminates with the target role group's development of a particular form of (coaching, school leadership, or facilitation) practice.
- The demonstrated means of supporting that learning process.

In our view, it is reasonable to extrapolate from the teacher learning literature because there is little research on supporting coaches', school leaders', and professional development facilitators' learning.

CURRENT LIMITATIONS OF DESIGN STUDIES

To this point, we have focused on the potential contributions of design studies. It is also important to take a critical stance by considering general limitations of such studies. In discussing classroom design studies, we highlighted a major current limitation that applies to the design research methodology more generally: the lack of an explicit, agreed-upon argumentative grammar. As we indicated, this is a severe weakness that must be addressed if design research is to become a mature methodology with explicitly codified standards that can be used to judge the quality of proposals for and reports of particular classroom and professional development design studies.

A second major limitation of both types of design studies concerns the limited attention that is typically given to issues of equity. It is important to acknowledge that the complexity of students' and teachers' learning, and of the designed learning environments, makes it impossible to specify completely everything that transpires in the course of a design study (Cobb, Confrey, et al., 2003). Choices therefore have to be made when framing a design study as a paradigmatic case of a broader class of phenomena. It is nonetheless striking that few classroom and professional development design studies have been conducted that focus explicitly on equity in student learning opportunities. In this regard, a classroom design conducted by Enyedy and Mukhopadhyay (2007) is a rare exception: they attempted to support students in making increasingly sophisticated statistical arguments by drawing on their out-of-school knowledge of city neighborhoods. In reporting their findings, Enyedy and Mukhopadhyay observed that there was an inherent tension between honoring the students' local knowledge while establishing disciplinary norms of argumentation.

The failure to attend explicitly to issues of equity reflects the assumptions that designs that effectively support all students' learning can be developed without attending explicitly to issues of equity, and that designs and forms of instructional practice that are judged to be productive are necessarily equitable. In our view, both these assumptions are suspect. In our view, attending to issues of equity in classroom design studies entails documenting the distribution of students' learning opportunities and, perhaps, the development of their mathematical identities (cf. Cobb, Gresalfi, & Hodge, 2009). In the case of professional development studies, it entails specifying intended instructional practices for participating teachers for which there is evidence that these practices will support the learning of diverse groups of students or, at a minimum, a conceptual analysis that indicates they have this potential.

A third major weakness concerns the frequent failure of researchers conducting classroom and PD design studies to design for scale when preparing for studies, thereby limiting the potential pragmatic payoff and relevance of their work beyond the research community. This weakness is especially evident in many classroom design studies, as researchers conducting such studies often give little consideration to the knowledge and skill that teachers would have to develop to enact the design effectively. In many cases, the learning demands appear to be unrealistic for most teachers given their current instructional practices. This weakness is also evident in many PD design studies, though in a less extreme form because researchers conducting this type of study cannot typically insulate teachers from the school settings in which they work. However, members of the research team often "camp out" in a small number of schools as central providers of support, failing to take account of the atypical expertise they bring to supporting the participating teachers' learning. One of the strengths of the design research methodology is that it enables researchers to explore what is possible in students' or teachers' learning. As a consequence, there is typically a significant discontinuity between typical forms of education and those that are the focus of classroom and PD design studies. However, the possibility that the design developed and refined in the course of a design study might contribute to improvements in classroom teaching and learning on a large scale will be significantly reduced unless researchers consider not merely their own but others' capacity to support students' or teachers' learning when formulating the design.

One approach for circumventing this limitation is to give at least as much weight to the problems of practice that school personnel identify as to researchers' assessments of what counts as theoretically interesting problems about students' or teachers' learning (cf. Bannan-Ritland, 2008). In this approach, researchers might take practitioners' concerns as their starting point and negotiate how those issues are framed so that the study is both pragmatically and theoretically significant. For example, a study that begins with teachers' concern about motivating students might reframe the focal issue in terms of cultivating students' mathematical interests or supporting their development of productive mathematical identities. In this and similar instances, the design research methodology would approach its full potential by exploring what is possible in students' or teachers' learning in a manner that is likely to have implications for educational improvement more generally.

The fourth limitation that we identified is specific to classroom design studies and concerns the lack of attention to the instructional practices of the teacher in the study. Most researchers who conduct classroom design studies would readily acknowledge that the study teacher plays a central role in supporting the participating students' learning. However, these teachers' instructional practices are rarely the focus of explicit analysis. This is unfortunate because these teachers typically enact relatively sophisticated practices. Analysis of their practice could therefore contribute to the delineation of key aspects of inquiry-oriented mathematics instruction. This would in turn help clarify the goals for teachers' learning that should be targeted in teacher professional development in general, and in professional development design studies in particular. Kwon et al.'s (2013) investigation of the development of mathematical argumentation is a rare exception as they framed the study teacher's discursive moves as a key designed support for the participating students' learning. Such attention to the study teacher's practices can inform efforts to support the implementation of instructional sequences developed during a classroom design study in other contexts.

The final limitations are specific to PD design studies. Most of the small number of studies of this type that have been conducted have focused exclusively on teachers' participation in the PD sessions. In our view, it is also essential to document changes in the quality of the participating teachers' instructional practices by, at a minimum, conducting pre- and post-observations of their classroom teaching. In addition, researchers conducting PD design studies have rarely attempted to document the settings in which the participating teachers work. As a consequence, the teacher group is, in effect, located in an institutional vacuum. The generalizability of study findings is thus threatened, making it difficult if not impossible for other researchers to adapt the PD design to the school contexts in which they are working. In reviewing reports of professional development design studies, it also became apparent that clear standards should be established for reporting such studies. A significant proportion of the reports we reviewed failed to provide information about the design principles that underpin the PD design, the conjectures about teachers' learning and the specific means of supporting their learning, the relation between teachers' participation in professional development activities and their classroom practice, and the aspects of school settings that mediate teachers' development of the intended forms of instructional practice. Carefully planned and executed PD design studies that are adequately reported are urgently needed as they can make critical contributions to the development and refinement of practice-specific professional development theories. The failure to include this essential information in published reports makes it difficult for research teams to learn from and build on prior studies. This, in turn, limits the possibility that the field will develop robust theories regarding how teachers can be supported to develop productive forms of practice.

REFERENCES

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In G. Sykes & L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3–32). San Francisco: Jossey Bass.

Ball, D. L., Sleep, L., Boerst, T., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *Elementary School Journal, 109*(5), 458–474.

Bannan-Ritland, B. (2008). Teacher design research: An emerging paradigm for teachers' professional development. In A. E. Kelly, R. Lesh & J. Y. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering and mathematics learning and teaching* (pp. 246–262). New York: Routledge.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2009). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education, 24*, 417–436.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*, 141–178.

Bruner, J. (1996). *The culture of education*. Cambridge, MA: Harvard University Press.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppesco, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.

Byrnes, J. P. (1992). The conceptual basis of procedural learning. *Cognitive Development, 7*(2), 235–257.

Cheeseman, J., McDonough, A., & Ferguson, S. (2012). The effects of creating rich learning environments for children to measure mass. In J. Dindyal, L. P. Cheng & S. F. Ng (Eds.), *Mathematics education: Expanding horizons (Proceedings of the 35th annual conference of the Mathematics Education Research Group of Australasia)*. Singapore: MERGA.

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*, 81–89.

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

Cobb, P., Gresalfi, M., & Hodge, L. L. (2009). An interpretive scheme for analyzing the identities that students develop in mathematics classrooms. *Journal for Research in Mathematics Education, 40*, 40–68.

Cobb, P., & Jackson, K. (2012). Analyzing educational policies: A learning design perspective. *The Journal of the Learning Sciences, 21*(4), 487–521.

Cobb, P., McClain, K., Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher, 32*(6), 13–24.

Cobb, P., & Steffe, L. (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education, 14*, 83–94.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *The Journal of the Learning Sciences, 10*(1–2), 113–163.

Cobb, P., & Yackel, E. (1998). A constructivist perspective on the culture of the mathematics classroom. In F. Seeger, J. Voigt & U. Waschescio (Eds.), *The culture of the mathematics classroom: Analysis and changes* (pp. 158–190). New York: Cambridge University Press.

Cobb, P., Yackel, E., & Wood, T. (1995). The classroom teaching experiment. In P. Cobb & H. Bauersfeld (Eds.), *Emergence of mathematical meaning: Interaction in classroom cultures* (pp. 17–24). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cobb, P., Zhao, Q., & Dean, C. (2009). Conducting design experiments to support teachers' learning: A reflection from the field. *Journal of the Learning Sciences, 18*, 165–199.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12.

Collins, A. (1992). Toward a design science of education. In T. Scanon & T. O'Shey (Eds.), *New directions in educational technology* (pp. 15–22). New York: Springer.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad.* Dallas, TX: National Staff Development Council.

DeCorte, E., Greer, B., & Verschaffel, L. (1996). Mathematics learning and teaching. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 491–549). New York: Macmillan.

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

diSessa, A. A. (1992). Images of learning. In E. d. Corte, M. C. Linn, H. Mandl & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 19–40). Berlin: Springer.

diSessa, A. A. (2002). Students' criteria for representational adequacy. In K. Gravemeijer, R. Lehrer, B. v. Oers & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 105–129). Dortrecht: Kluwer.

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*(3), 293–331.

diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences, 13*, 77–103.

Doorman, M., Drijvers, P., Gravemeijer, K., Boon, P., & Reed, H. (2013). Design research in mathematics education: The case of an ict-rich learning arrangement for the concept of function. In T. Plomp & N. Nieveen (Eds.), *Educational design research—Part B: Illustrative cases* (pp. 425–446). Enschede, The Netherlands: SLO.

Enyedy, N., & Mukhopadhyay, S. (2007). They don't show nothing I didn't know: Emergent tensions between culturally relevant pedagogy and mathematics pedagogy. *Journal of the Learning Sciences, 16*, 139–174.

Fishman, B., Marx, R. W., Blumenfeld, P., & Krajcik, J. S. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences, 13*, 43–76.

Forman, E. A. (2003). A sociocultural approach to mathematics reform: Speaking, inscribing, and doing mathematics within communities of practice. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 333–352). Reston, VA: National Council of Teachers of Mathematics.

Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal, 38*(3), 653–689.

Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225–256). Greenwich, CT: Information Age Publishers.

Freudenthal, H. (1973). *Mathematics as an educational task.* Dordrecht, The Netherlands: Reidel.

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*, 915–945.

Gravemeijer, K. (1994a). *Developing realistic mathematics education*. Utrecht, The Netherlands: CD-ß Press.

Gravemeijer, K. (1994b). Educational development and developmental research. *Journal for Research in Mathematics Education, 25*, 443–471.

Gravemeijer, K., & Cobb, P. (2006). Design research from the learning design perspective. In J. v. d. Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 17–51). London: Routledge.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P.W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*(9), 2055–2100.

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and teaching: Theory and practice, 15*(2), 273–289.

Grossman, P., O'Keefe, J., Kantor, T., & Delgado, P.C. (2013, April). *Seeking coherence: Organizational capacity for professional development targeting core practices in English language arts*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Hall, R. (2001). Schedules of practical work for the analysis of case studies of learning and development. *Journal of the Learning Sciences, 10*, 203–222.

Henningsen, M., & Stein, M.K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524–549.

Hesse, M.B. (1965). *Forces and fields: A study of action at a distance in the history of physics*. Totowa, NJ: Littlefield, Adams. & Co.

Hill, H.C., Ball, D.L., & Schilling, S.G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*(4), 372–400.

Hill, H.C., Schilling, S.G., & Ball, D.L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal, 105*(1), 11–30.

Horn, I.S. (2005). Learning on the job: A situated account of teacher learning in high school mathematics departments. *Cognition and Instruction, 23*, 207–236.

Hoyles, C., Noss, R., & Pozzi, S. (2001). Proportional reasoning in nursing practice. *Journal for Research in Mathematics Education, 32*(1), 4–27.

Jackson, K., & Gibbons, L. (2014, April). *Accounting for how practitioners frame a common problem of practice—students' struggle in mathematics*. Paper presented at the National Council of Teachers of Mathematics Research Conference, New Orleans, LA.

Kazemi, E., & Franke, M.L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education, 7*, 203–235.

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development: Attending to the coevolution of teachers' participation across contexts. *Journal of Teacher Education, 59*, 428–441.

Kelly, A.E. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences, 13*, 115–128.

Kwon, O.N., Ju, M-K., Kim, R.Y., Park, J.H., & Park, J.S. (2013). Design research as an inquiry into students' argumentation and justification: Focusing on the design of intervention. In T. Plomp & N. Nieveen (Eds.), *Educational design research—Part B: Illustrative cases*. Enschede, The Netherlands: SLO.

Lamberg, T., & Middleton, J.A. (2009). Design research perspectives on transitioning from individual microgenetic interviews to a whole-class teaching experiment. *Educational Researcher, 38*(4), 233–245.

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., & Franke, M.L. (2010). Using designed instructional activities to enable novices to manage ambitious mathematics teaching. In M.K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 129–141). New York: Springer.

Lampert, M., Franke, M.L., Kazemi, E., Ghousseini, H., Chan Turrou, A., Beasley, H., . . . Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education, 64*(3), 226–243.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. London: Cambridge University Press.

Lehrer, R., & Kim, M.J. (2009). Structuring variability by negotiating its measure. *Mathematics Education Research Journal, 21*(2), 116–133.

Lehrer, R., Kim, M.J., & Jones, S. (2011). Developing conceptions of statistics by designing measures of distribution. *ZDM: The International Journal on Mathematics Education, 43*, 723–736.

Lehrer, R., Schauble, L., Strom, D., & Pligge, M. (2001). Similarity of form and substance: Modeling material kind. In S. M. Carver & D. Klahr (Eds.), *Cognition and Instruction: Twenty-five years of progress* (pp. 39–74). Mahwah, NJ: Lawrence Erlbaum Associates.

Lesh, R., & Kelly, A. E. (1997). Teachers' evolving conceptions of one-to-one tutoring: A three-tiered teaching experiment. *Journal for Research in Mathematics Education, 28*(4), 398–430.

Little, J. W. (2003). Inside teacher community: Representations of classroom practice. *Teachers College Record, 105*, 913–945.

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32*(1), 17–20.

Lobato, J. (2012). The actor-oriented transfer perspective and its contributions to educational research and practice. *Educational Psychologist, 47*(3), 232–247.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3–11.

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education, 64*(5), 378–386.

Menchinskaya, N. A. (1969). Fifty years of Soviet instructional psychology. In J. Kilpatrick & I. Wirzup (Eds.), *Soviet studies in the psychology of learning and teaching mathematics* (Vol. 1, pp. 5–27). Stanford, CA: School Mathematics Study Group.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4–15.

Saldanha, L., & Thompson, P. (2007). Exploring connections between sampling distributions and statistical inference: An analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education, 2*, 270–297.

Schoenfeld, A. H. (2006). Design experiments. In P. B. Ellmore, G. Camilli, & J. Green (Eds.), *Complementary methods for research in education*. Washington, DC: American Educational Research Association.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of video club. *Teaching and Teacher Education, 20*, 163–183.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*, 114–145.

Simon, M. A., Saldanha, L., McClintock, E., Karagoz Akar, G., Watanabe, T., & Ozgur Zembat, I. (2010). A developing approach to studying students' learning through their mathematical activity. *Cognition and Instruction, 28*, 70–112.

Simpson, G., Hoyles, C., & Noss, R. (2006). Exploring the mathematics of motion through construction and collaboration. *Journal of Computer Assisted Learning, 22*(2), 114–136.

Smith, J. P., & Thompson, P. W. (2007). Quantitative reasoning and the development of algebraic reasoning. In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 95–132). Mahwah, NJ: Lawrence Erlbaum Associates.

Steffe, L. (1983). The teaching experiment methodology in a constructivist research program. In M. Zweng, T. Green, J. Kilpatrick, H. Pollak, & M. Suydam (Eds.), *Proceedings of the Fourth International Congress on Mathematical Education* (pp. 469–471). Boston: Birhauser.

Steffe, L., & Kieren, T. E. (1994). Radical constructivism and mathematics education. *Journal for Research in Mathematics Education, 25*, 711–733.

Steffe, L., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 267–307). Mahwah, NJ: Lawrence Erlbaum Associates.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313–340.

Stephan, M., & Akyuz, D. (2012). A proposed instructional theory for integer addition and subtraction. *Journal for Research in Mathematics Education, 43*, 428–464.

Streefland, L. (1991). *Fractions in realistic mathematics education. A paradigm of developmental research.* Dordrecht, The Netherlands: Kluwer.

Taylor, S. J., & Bogdan, R. (1984). *Introduction to qualitative research methods: The search for meanings* (2nd ed.). New York: John Wiley & Sons.

Thompson, P. W. (1994). Images of rate and operational understanding of the Fundamental Theorem of Calculus. *Educational Studies in Mathematics, 26*, 229–274.

Thompson, P. W. (1996). Imagery and the development of mathematical reasoning. In L. P. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 267–283). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thompson, P. W., & Saldanha, L. A. (2000). Epistemological analyses of mathematical ideas: A research methodology. In M. Fernandez (Ed.), *Proceedings of the Twenty-Second Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 403–407). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Toulmin, S. E. (1958). *The uses of argument.* Cambridge: Cambridge University Press.

Treffers, A. (1987). *Three dimensions: A model of goal and theory description in mathematics instruction—The Wiskobas Project.* Dordrecht, The Netherlands: Reidel.

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education, 24*, 244–276.

Wawro, M., Rasmussen, C., Zandieh, M., & Larson, C. (2013). Design research within undergraduate mathematics education: An example from introductory linear algebra. In T. Plomp & N. Nieveen (Eds.), *Educational design research—Part B: Illustrative cases* (pp. 905–925). Enschede, The Netherlands: SLO.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* New York: Cambridge University Press.

Zawojewski, J., Chamberlin, M., Hjalmarson, M. A., & Lewis, C. (2008). Developing design studies in mathematics education professional development: Studying teachers' interpretive systems. In A. E. Kelly, R. Lesh, & J. Y. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering and mathematics learning and teaching* (pp. 219–245). New York: Routledge.