Part 1

Introduction
1 Enabling Innovations in Education and Systematizing their Impact

Anthony E. Kelly
George Mason University

John Y. Baek
Center for Advancement of Informal Science Education

Richard A. Lesh
Indiana University

Brenda Bannan-Ritland
George Mason University

Introduction
Design research in education is directed at developing, testing, implementing, and diffusing innovative practices to move the socially constructed forms of teaching and learning from malfunction to function or from function to excellence (Kelly, 2003). In this way, the open character of design together with the self-imposed constraints of research are placed in service of the growing demands of society for citizens and knowledge workers who can add to and take advantage of emerging technologies and the explosion in scientific knowledge.

Although some policy-makers view the need for improved human capital as primarily a vehicle for supporting national competitiveness (Committee on Science, Engineering, and Public Policy, 2007; Domestic Policy Council, 2006), we view the development of talent among learners and productivity among teachers as a universal good, in concert with a motto attributed to President John F. Kennedy that, “a rising tide raises all boats.” For this reason and because learning challenges are global, this book draws on both US and international contributions.

The preface of this book describes the burgeoning growth in scientific knowledge and the growing complexity of science as a practice, globally. How should educational researchers respond to these challenges so as to reengineer models of schooling that are little changed since the Industrial Revolution?

A central question for educational research is how to design interventions that move beyond describing “what is?” or confirming “what works?” to designing “what strategy or intervention might work better?”—especially in content domains with little prior research (Cobb et al., 2003). Or, more expansively, “what systemic reengineering of learning environments might work better to teach students and teachers to respond to the opportunities rapidly unfolding in modern science?”

Modern educational interventions must respond to new scientific knowledge emerging from technology-infused, internet-intensive, highly social, networked science. Given the pace of change, many educational researchers find inadequate guidance and theorizing from psychology or cognitive science or from the findings about the effectiveness of prior (but perhaps dated) interventions (e.g., Brown, 1992). As
McCandliss et al. (2003) argue, a collaborative dialogue among methods is needed that yields both internally and externally valid findings. And, we would add, a dialogue that responds to local resource demands and capabilities, including human creative capital.

We believe that the hybrid methodology now known as educational design research, which was initiated in education research circles in the early 1990s (e.g., Brown, 1992; Kelly, 2003) offers significant promise in this regard. Design research is characterized in this book, but other sources should be consulted for a more complete treatment (e.g., van den Akker Gravemeijer et al., 2006; Barab & Squire, 2004; Kelly, 2003, 2004, 2007). In this chapter, we wish to explicate design research by: (a) comparing its commitments to those of more traditional approaches, and (b) by illuminating the research themes across the chapters in this volume.

**Commissive Spaces in Research Practice**

First, we contrast the commissive spaces of more systematized research approaches to that of the newer design research methodology. Commissive spaces describe the background commitments, rules, and shared implicit and explicit values held by members of a community that permit and undergird the conversations and actions that define that community (Design-Based Research EPSS, 2006; Kelly, 2006).

To illustrate the commissive space of traditional research, examine the orientation for framing of educational research that was outlined by the National Research Council:

> To simplify matters, the committee recognized that a great number of education research questions fall into three (interrelated) types: description—What is happening? cause—Is there a systematic effect? and process or mechanism—Why or how is it happening? (2002: 99)

The commissive space within which these questions are posed and answered assumes (for those editors, at least) commitments to evidence, warrant, and methods that privilege the stance of the researcher-as-observer of some more-or-less remote phenomenon. The goal of answering these questions is to advance a field of knowledge by adding reliable observations, facts, principles, or descriptions of mechanisms to an existing and growing body of laws, regularities and mechanisms. The methods for achieving that goal (particularly for “causal” questions) include those such as randomized clinical trials, which value commitments to: (a) an a priori alignment with theory (if available), (b) rigorous elimination of competing explanations via random assignment to condition, (c) the identification and implementation of a known and describable “treatment” implemented with some desirable “fidelity,” and (d) commitments to standardized (i.e., pre-existing) measures of achievement (see also, Coalition for Evidence-Based Policy, 2003). The immediate audience for such research findings is peer researchers, and to the extent that the research is “applied,” the audience may extend to various interpreters, implementers, or users of the research, including policy-makers or practitioners.

Actors in this commissive space favor: (a) a convergence of observation, methods, and metrics with a priori stances, (b) a tendency not to pursue what appear to be tangential or emergent phenomena, (c) a proclivity to devalue context, and (d) a commitment to valuing the researcher’s assumed objective stance over the subjective stance of the “subjects.”
By contrast, while the design research commissive space shares many of the commitments of more traditional approaches, design researchers foreground the fluid, empathetic, dynamic, environment-responsive, future-oriented and solution-focused nature of design. Design researchers often recruit the creativity of students, teachers or policy-makers not only in prototyping solutions, but also in enacting and implementing the innovation, and in documenting the constraints, complexities, and trade-offs that mold the behavior of innovative solutions in contexts for learning. By observing and participating in the struggles of the design, and the implementation or diffusion of an innovation, design researchers may learn not only how to improve an innovation, but also how to conduct just-in-time theory generation and testing within the context of design processes and in the service of the learning and teaching of content.

Indeed, design researchers posit that more than a single theory may be required to describe, explain or predict the success or failure of an innovation (e.g., Lesh & Hjalmarson, this volume). On the other hand, in some cases, theoretical or modeling work will fail to provide an adequate account of the success or failure of an innovation, and the designer and users will “satisfice” and “work with” procedures and products that appear to be advancing shared goals even in the absence of final understanding (e.g., Bannan-Ritland & Baek, this volume; Barab, Baek, Schatz, Scheckler, & Moore, this volume). Such a result is often the case in medical practice in which drugs or other therapies have successful outcomes for which mechanistic explanations are currently unavailable. Moreover, innovations are sometimes suggested and revised on the basis of poorly articulated theoretical groundings, championed nonetheless by powerful stakeholders (e.g., Bannan-Ritland & Baek, this volume; Brazer & Keller, this volume).

Design researchers also embrace, in some form, the traditional instructional design ADDIE tradition (Analysis, Design, Development, Implementation, and Evaluation). However, they move beyond instructional design as craft knowledge toward understanding the know-how/know-why of the design (e.g., Clements, this volume). Since design researchers are co-designing the environments they are studying, they can gain insight into who might best deploy solutions or be likely to adopt them in practice (e.g., Ejersbo, Engelhardt, Frølunde, Hanghøj, Magnussen, & Misfeldt, this volume). They may also dynamically articulate what systemic and contextual resources are present or lacking (e.g., Cobb & Gravemeijer, this volume). It should be noted, however, that the principles of ADDIE still hold significant value for design researchers (see Barab et al., this volume). Moreover, the ADDIE model, itself, has become more complex over time (Clark, 2007).

Thus, design research inhabits and defines a commissive space of novelty tempered by research evidence that unfolds over multiple cycles within a study or across a program of studies (Bannan-Ritland, 2003), drawing, sometimes opportunistically from—and organically adding to—many sources of information and knowledge. Design research attempts to: (a) help design innovations, (b) explain their effectiveness or ineffectiveness, theoretically, and (c) re-engineer them where possible, while adding to the science of design itself (e.g., Bannan-Ritland & Baek, this volume; Barab et al., this volume; Cobb & Gravemeijer, this volume; Kali, this volume; Martinez et al., this volume; Roschelle, Tatar, & Kaput, this volume). Design research strives to mine the massive “information loss” that can occur when interventions are designed “off-line” and applied with little revision, and with little attention to the causes of failure or success. Within this larger frame of design processes, design research draws, when appropriate, from the strengths of other research traditions, and multiple approaches to assessment.
We want to stress that design research methods are a work in progress and are not advocated as a panacea for promoting and developing educational innovations. As has been stressed from early work in this field (e.g., Bannan-Ritland, 2003; Clements, this volume), design research methods make sense at some points in a larger developmental process but not at others.

The Integrative Learning Design Framework (ILDF: Bannan-Ritland, 2003; Bannan-Ritland & Baek, this volume) provides one overarching structure within which to position design problems, selected methods, and design moves in the design research process. By acknowledging the fluid, dynamic, evolving nature of design in design research and the many layers of decisions and judgments required in design (cf. Brazer & Keller, this volume), the ILDF encourages design researchers to consider other factors beyond theoretical conjectures. The generation and refinement of theory, the selection of research questions, methods, and design features become integrated with pragmatic processes of design (see also, Barab et al., this volume).

As a second example, Clements (this volume) describes the Curriculum Research Framework (CRF), which consists of ten phases of the development research process that warrant claiming that a curriculum is based on research. While Clement sees a role for design studies, he recognizes that other research and development strategies are necessary to meet the goals of a complete curriculum research and development program. As has been noted, elsewhere (e.g., Kelly, 2003, 2004, 2006), design studies are unable to control many variables in complex settings; rarely analyze in full the large amount of data collected before the next cycle of revision, enactment, and analysis takes place; and may value different data from the perspectives of different participants. As a result the paths and products may be arbitrary to an extent and generalization may be difficult.

Clements argues that randomized trial experimental designs provide some of what design studies cannot. In addition, he notes that the use of phases in the a priori foundations and learning model categories of the CRF provide useful constraints and theoretical groundings for design experiments. Conversely, design experiments, as well as other methods such as teaching experiments and classroom-based teaching experiments, can help accomplish what randomized trials cannot (Kelly & Lesh, 2000). These methods include conceptual and relational, or semantic, analyses, and are theoretically grounded. They allow researchers to build models of learning and of teaching interactions. Ultimately, as Clement elucidates, because it includes a coherent complement of methods, the CRF has built-in checks and balances that address the limitations of each method, with concentration on the learning model especially useful for maintaining theoretical and scientific foci.

As a third example, for Middleton et al. (this volume) design research is analogous to engineering design. Among the necessary conditions for calling aspects of product development “design research” are the provision of a coherent model with an explicit chain of reasoning connecting the features of the innovation and the social and cognitive systems they purport to affect, and the development of methods and generation of evidence that have the potential to replicate and transport across studies and implementation settings. The design cycle for Middleton et al. draws on Bannan-Ritland (2003), and is presented as having seven distinct stages, beginning at conceptualization of the problem to be addressed and extending to market and revision of an actual product. These stages consist of: (a) grounded models, (b) development of artifact, (c) feasibility studies, (d) prototyping and trialing, (e) field study, (f) definitive test, and...
Factors in Current Design Research in Educational Innovations

In the prior section we described how the commissive space that design research occupies differs from traditional research approaches. In this section we continue to explore the commissive space by highlighting what we consider to be the important themes in the current design research reported by the contributing authors. In providing this commentary of the work, we intend to illuminate some underlying factors that are necessary for designing and researching educational innovations. The core idea that provides the most resonance in the design research literature is the idea of iteration, the capacity and knowledge to modify the intervention when it appears not to work or could be improved. We have outlined below some of the bases on which design researchers decide to launch an iteration of some innovation. Some design researchers use a theoretical base, some use insights from measurement or instrumentation, others use input from team members or stakeholders and review combinations of these sources.

Iteration and Assessment

Assessment is a critical problem for design researchers since modeling learning during re-design cycles defines the approach: whether learning by students (e.g., Lobato, this volume; Cobb & Gravemeijer, this volume; Rasmussen & Stephan, this volume), teachers (Bannan-Ritland, this volume; Zawojewski, Chamberlin, Hjalmarson, & Lewis, this volume), policy-makers (Brazer & Keller, this volume), or researchers (e.g., Bannan-Ritland & Baek, this volume; Lesh, Kelly, & Yoon, this volume).

In design research as currently practiced, assessment is not directed at some summative sense of learning, though a summative measure of student learning would be central to later attempts at confirmatory studies, i.e. to show local impact (Bannan-Ritland, 2003). That being said, Hake (this volume)—arguing that physics education research is a genre of design research—uses gain scores in pre-post testing to validate his process model of causality (Maxwell, 2004). For him, assessment acts as a guide for improving the teaching and learning of physics, in a broad programmatic sense.

Design research also differs from formative assessment with regard to the student’s knowledge end state and how feedback loops are enacted. Formative assessment is the gathering of data relative to some predetermined fixed point, providing feedback that informs the students and teacher of their current knowledge state in relation to some end state (see Black & Williams, 1998). In design research, assessment may be used formatively in order to dynamically determine progress toward mastery of disciplinary knowledge (e.g., Cobb & Gravemeijer, this volume) or to guide the design of a prototype and to inform its iterative re-design as necessary or both. In fact, sensitivity to assessment practices themselves may inform changes to the act of assessment itself (e.g., Lobato, this volume; Lesh et al., this volume). Ultimately, design researchers are challenging the assumptions about learning, teaching, and knowing that underlie available assessment techniques, not only in terms of the psychometric assumptions (like item response theory), but also the function of assessment itself within and across the stages of design research (see Sloane & Kelly, this volume).
Assessment and Theory Building

Within design research circles, there is a debate about the status of theory generation and the centrality of theorizing in the design process (e.g., see Bannan-Ritland & Baek, this volume; Barab et al., this volume; Design-Based Research Collaborative, 2003; Middleton, Gorard, Taylor, & Bannan-Ritland, this volume). Theory acts in two different ways in design research: the theory of learning that is assumed by the researcher, and generation of theory through the process of design. Where do the theories used in design research projects come from? Though design researchers see the world through participant eyes, they are informed by relevant theory, pragmatically selecting the most relevant at the time. Theories may draw from studies of cognition, perception, motivation, neuropsychological function, interpersonal communication, or, more generally, from principles of social organization.

Decision points are not always informed by learner assessment. Pragmatic, political, and participatory factors can influence the decision to iterate. Being responsive to externalities, these factors emerge within a design context that is “ready” to listen and address them directly into the design (e.g., Wolf & Le Vasan, this volume).

Assessment and Teaching Experiments

To the extent that the assessments of learning-in-progress are explicit and rigorous and tied to disciplinary knowledge, they can also inform local instructional theory building (Cobb & Gravemeijer, this volume). The approach known as teaching experiments, particularly in mathematics education, have been documented elsewhere (e.g., Steffe & Thompson, 2000). More recently, teaching experiments have been reconceptualized within the more general framework of design research (Cobb & Gravemeijer, this volume; Cobb et al., 2003).

Cobb and Gravemeijer’s chapter differentiates among three phases of a design experiment: (a) preparing for the experiment, (b) experimenting in the classroom to support students’ learning, and (c) conducting retrospective analyses. The initial preparation phase is crucial to the success of the experiment and can be extensive, especially when there is little prior research on which to build. Key issues that need to be addressed in this phase include clarifying instructional goals by identifying the central ideas in a particular domain and documenting the instructional starting points both by drawing on the relevant literature and by conducting initial assessments as part of the pilot work. In addition, it is essential to delineate an envisioned learning trajectory that consists of conjectures about both a learning process that culminates with the prospective instructional goals and the specific means of supporting that learning process. As they illustrate, the primary objective in the second phase of experimenting in the classroom is not to demonstrate that the envisioned learning trajectory works but to improve the envisioned trajectory by testing and revising conjectures about both the prospective learning process and the specific means of supporting it. This conjecture-testing process depends crucially both on the kinds of data that are collected in the course of an experiment, and on the adequacy of the analytic framework that is used to document both students’ learning, and the evolution of the classroom learning environment. The intent of retrospective analyses conducted during the third phase of an experiment is to contribute to the development of a domain-specific instructional theory by documenting both a substantiated learning trajectory and the demonstrated means of supporting students’ learning along that trajectory. This theoretical goal gives rise to a number of methodological challenges, not the least of which is to differentiate between aspects of
the classroom learning environment that are necessary and those that are merely contingent in supporting student learning. The additional methodological issues that the authors consider include the trustworthiness, repeatability, and generalizability of the findings.

Assessment and Variation Theory

In this book, there are some attempts to tie assessment practices to some larger theory. For example, Holmqvist, Gustavsson, and Wernberg (this volume) show how student performance on brief, objective tests tied directly to a narrow set of concepts can guide the redesign of instruction, guided by variation theory, which grows from perceptual psychology (Gibson, 1986) and Bransford and Schwartz’s (1999) Preparation for Future Learning framework. According to variation theory, learning occurs as a function of how attention is selectively drawn to critical aspects of the object of learning by the student. Teachers, as a function of their competence, may direct attention toward critical aspects (e.g., by using contrasting cases) or away from critical aspects and toward irrelevant aspects, sometimes to the students’ detriment. For example, to learn the definition of the size of angles, the student should focus on the amount of rotation between the two rays defining the angle, rather than on the length of the rays themselves. A study guided by variation theory would, for example, examine both the teaching and the student learning of angles in order to determine if by selective attention or teacher (mis)direction whether the student was focusing on critical, irrelevant, or inaccurate aspects of the definition of an angle. Micro-analyses of classroom activities and of test-item performance would identify any misgeneralizations in response to teaching episodes, which would guide the re-design of instructional strategy or content (e.g., designing a lesson that showed that angle size was not dependent on the length of rays).

Assessment and Transfer

The chapter by Lobato examines design cycles in educational settings from the perspective of transfer of learning. In the traditional view, a learner is assumed to learn from one task, and apply (or “transfer”) the insight to a second task that is similar (from the researcher’s perspective, at least). The study of transfer of learning is described in detail in this chapter, in which the author outlines the major competing views (including Bransford and Schwartz’s). According to Lobato’s (2003) own view (that of actor-oriented transfer), the judgment of the “similarity” of the tasks must take into account the interpretations of both tasks by the learner, which may (from the traditional researcher’s perspective) be viewed as non-normative or lead to incorrect performance.

As the actor-oriented perspective has matured, Lobato has tackled the notion of “mechanism” or more accurately how the generalization of learning experiences is brought about as a socially-situated phenomenon. She and her research team are investigating how the ways in which students generalize their learning experiences are related conceptually to the particular mathematical regularities that students come to notice in mathematics classrooms when multiple sources of information compete for their attention. The “focusing interactions” framework that they have developed allows them to document the social organization of this “noticing” behavior and coordinate it with psychological processes. The significance of this work for mathematics education researchers and teachers is found in the demonstration that the durable concepts that students generalize from instruction are influenced, not simply by the macro-level
actions often recommended by reforms (e.g., use of collaborative groups, inquiry-oriented materials, or manipulatives) but also by many subtle micro-features of instruction that come into play as different mathematical foci emerge in classrooms.

Lobato’s chapter raises important questions for educational practice in general, particularly questions of construct validity in the design of assessments of learning, both proximal and distal. Like Minstrell (e.g., Minstrell, 1992; van Zee & Minstrell, 1997), Lobato challenges educational researchers to consider the partial understanding of learners (given the learner’s perspective and level of expertise) as part of a foundation upon which later learning may be built. Lobato’s work raises questions not just about similarity of assessment tasks, but argues for a more forgiving scoring rubric, closer to Minstrell’s “facets of understanding” model (see also the work of diSessa and colleagues: Cobb et al., 2003; diSessa, 1993; diSessa & Cobb, 2004).

Lobato’s model also implicates features of instruction in the learner’s (mis)conceptions about topics. Teachers may inadvertently draw attention to irrelevant features of the task or mislead learners by misconstruing central conceptual aspects during a lesson. Lobato terms this occurrence a “focusing phenomenon,” which she describes in her chapter. The reader may wish to compare Lobato’s treatment of focusing phenomena with the Holmqvist et al.’s (this volume) use of “variance theory” to inform re-design of lessons in their work.

IMPLICATIONS FOR DESIGN RESEARCH

Lobato’s work has a number of implications for design research (which she began to spell out in Lobato, 2003). Her work challenges “humble theory” building in design research to identify which of the views of transfer of learning is assumed to be dominant during the design process. Each model of transfer has different implications for who has the privileged position in designing the artifact and on whose knowledge base to decide when it is time to revise the design. More generally, the design of the assessments both formative and summative should similarly reflect a commitment to a model of transfer, since design research strives to support not just achievement, but also learning to learn in new contexts.

Thus, Lobato’s work has implications not only for early stage, prototyping, or discovery research, it also has implications for the design of assessments during more definitive testing (e.g., using randomized clinical trials) or later during diffusion stages of research (e.g., Bannan-Ritland, 2003).

Assessment and a Theory of Change

One goal of the design researcher is to build instructional artifacts that improve student learning. Sloane and Kelly (this volume) stress the importance of specifying a theory of learning (i.e., one based on changes in what a learner masters), rather than on a theory of achievement, which reduces learning to the attainment of an assumed terminal point. They encourage the use of models that can capture changes in learning over time that are qualitative, non-linear, or non-additive. They highlight some of the definitional features of change implicit in education research to guide design researchers as they move to quantify student growth over time. They note that the challenges posed to design researchers in modeling change afflict, equally, the modeling of change even by those with mastery of current statistical modeling formalisms.
Assessment and Learning to Learn

Given the emphasis on design processes in design research, some authors champion approaches to assessment that place the task of design squarely on the learners, be they students (e.g., Lesh et al., this volume; Roschelle et al., this volume) or teachers (e.g., Bannan-Ritland, this volume; Zawojewski et al., this volume). Lesh and his colleagues (e.g., Lesh & Lamon, 1992; Lesh et al. 2000) have argued for decades that assessment should fold back on the learner the responsibility for documenting growth toward and the construction of a solution. Thus, Lesh advocates the use of model-eliciting (or thought-revealing) problems, where the “answer” is a model or blueprint for solving a class of problems, not a single number. In Lesh et al. (this volume), the authors show how a similar approach to assessment can work in a multitiered fashion so that students’ growth in modeling provides data for teachers’ growth in modeling of student learning, which, in turn, provides data for researchers’ growth in modeling both teacher and student learning.

The goal of design research in Lesh’s model is to engage the participants in a study (including the researchers) in ongoing reflexive cycles of design and re-design, so that while there is an end-in-view for the study (and for learning), the constitution of that end-in-view and how progress toward constructing and measuring it is co-determined, primarily, by the learner. Stated differently, Lesh views assessment itself as a dynamic object of study and a primary driver within design research studies.

Assessment and Codifying Innovation or Efficiency

Schwartz, Chang, and Martin (this volume) are interested in the development of instrumentation as an effective way to help move design research from the construction of new possibilities to openness to tests of causality and generality. Their argument is that the creativity behind an innovation can be expressed via instrumentation, which then exists after the design research to support efforts at efficiency in implementation (on this point, see Martinez et al., this volume; Roschelle et al., this volume).

Similar to the approach suggested by Lesh and his colleagues, Schwartz et al. describe design research studies in which high school students are expected to invent a way to measure variability. They show that this demand for innovation later helps the students to learn efficient canonical solutions. For researchers, they demonstrate how developing a novel double-transfer instrument enabled the measurement of an important type of learning that conventional measures miss. By working on instrumentation, the researchers equip themselves with the tools necessary to replicate the experiment and measure their progress.

Finally, Schwartz et al. discuss the need for warrants of innovation by distinguishing between innovating knowledge and innovating practice. Both are goals of design studies, but individual instances of research often emphasize one or the other. One warrant for both types of innovation is based on the reconciliation of inconsistencies: a thing is recognizably new and innovative when it resolves a previous contradiction in knowledge or in social practices. This warrant relates to the idea of praxis, where a theory is proven by becoming a reality. They believe praxis is a highly relevant way of warranting educational research that seeks to create new ways for students to learn.

Both Schwartz and Lesh and their colleagues are concerned with developing in students the capacity to learn-to-learn, i.e. to learn to respond to challenges from outside the learning system. As our technological society continues to change and advance, new challenges and emergent phenomena will be the rule. Students should be adequately
prepared to face the chaos and complexity of the world rather than shy away from it (e.g., Saari, this volume). The transformative goals of learning and education are set in design research: not to teach reading, but to build readers; not to teach mathematics, but to build mathematicians. Metaphorically, learning how to be a good traveler is as important as learning how to reach and enjoy Paris.

Assessment, Argument, and Collective Activity

One of the challenges facing design researchers is how to assess collective activity in a large group in order to determine if there is “general understanding” of some concept. As we have seen, many researchers view learning at the level of the individual. Recently, some researchers have begun to posit the argument that if classroom discourse appears to indicate that some new construct has become generally accepted in the group, then it can be taken “as-if-shared” by every member in the group, but without independent confirmation by the researcher at the member level.

Rasmussen and Stephan (this volume) developed a three-phase methodological approach for documenting the collective activity of a classroom community. Most interesting, they place the concept of collective activity on a strong footing by basing their method on the systematic use of Toulmin’s argumentation scheme over extended classroom lessons. Toulmin’s basic model of argumentation describes the structure and function of an argument in terms of four parts: the data, the claim, the warrant, and the backing. They developed the following two criteria for determining when mathematical ideas function as if shared: (a) when the backings and/or warrants for an argumentation no longer appear in students’ explanations (i.e., they become implied rather than stated or called for explicitly, no member of the community challenges the argumentation, and/or if the argumentation is contested and the student’s challenge is rejected), and (b) when any of the four parts of an argument shifts position (i.e., function) within subsequent arguments and is unchallenged (or, if contested, challenges are rejected). The usefulness of the methodology is two-fold. First, it offers an empirically grounded basis for design researchers to revise instructional environments and curricular interventions. Second, it offers an innovative way to compare the quality of students’ learning opportunities across different enactments of the same intervention. They illustrate the methodology with an example from a first-grade class learning to measure and from a university course in differential equations. They conclude the chapter by discussing issues of the generalizability and the trustworthiness of the methodology.

Documentation

Documentation is the archiving and indexing of artifacts of the design research process that serves as a way of gathering evidence of the effects of design changes, and serves to inform re-design if changes to a prototype prove ineffective. The process of documentation embraces yet exceeds the function of assessment. Bannan-Ritland and Baek (this volume) identify the importance of documenting theoretical, opportunistic, political, and practical decisions. Documentation of the progress of design can not only make explicit the influence of factors that support and constrain effective practice, but may also serve as the basis for adding to a science of design (see also, Kali, this volume).

In their chapter, Barab et al., using a design narrative methodology, illuminate the challenges of designing a web-based community, the Inquiry Learning Forum (ILF). They highlight the design challenges and successes and advance some theoretical assertions in hopes that others may more fruitfully carry out their own design work. A
challenging part of doing educational research on design-based interventions is to characterize the complexity, fragility, messiness, and eventual solidity of the design. All too often, designers simply report the ready-made structures, thereby obscuring or “black-boxing” the trajectories through which design decisions are made. This is problematic in that much of the theory generation process necessary to move from design work to design-based research occurs through an examination of these situated processes. Therefore, in helping others to determine the local utility of the derived theoretical assertions, a core goal of this chapter is to lay open and problematize the completed design in a way that provides insight into the “making of” the design. This process involves not simply sharing the designed artifact, but providing rich descriptions of context, guiding and emerging theory, design features of the intervention, and the impact of these features on participation and learning. The characterization of the ILF is intended to have both local resonance with the data as well as more global significance.

Multitiered Models for Teacher Professional Development

The chapter by Zawojewski, Chamberlin, Hjalmarson, and Lewis explores the use of multitiered design studies to support mathematics teachers’ professional development while producing generalizable theory about the interpretive systems teachers use to teach. They describe the design of educational objects (e.g., an algebra lesson, a student problem-solving approach sheet) that support professional development while making explicit the interpretive systems of the teachers, professional development facilitators, and teacher education researchers who use them. Two examples of multitiered design studies are examined. In the first example, “student thinking sheets,” teachers use, analyze, and re-design problem-solving approach description sheets that capture students’ thinking about complex mathematical problems. In the second example, “lesson study,” teachers design, teach, analyze, re-design, and re-teach a lesson designed to help students identify and mathematically express a pattern.

In both examples, a designed artifact simultaneously supports professional development while making teachers’ and professional development designers’ interpretive systems visible for study. Teachers examine classroom situations and their own interpretive systems, and researchers and professional development designers examine teachers’ interpretive systems (as well as their own) in order to derive principles for professional development that may be transportable to other contexts. An auditable trail of artifacts from the professional development work is the primary data source, allowing researchers to focus on the actual performance of the professional work. The design study approach is argued to be well suited to situations in which participants are expected to grow and improve in different ways, rather than necessarily converge toward a particular standard. Further, the approach is argued to be well suited to situations in which teachers address problems that they themselves have chosen. The notion of design study is proposed as a means to embrace such diversity in systems where changes in one part of the system can reverberate throughout the system (e.g., teacher change leads to change in the professional development provider perspective, leads to change in the professional development materials and implementation, which influences students’ learning, which impacts teachers’ development, etc.).

Statistical and Other Modeling Formalisms

Along with theory building, design researchers also build instructional artifacts to support student learning. In the practice of design research, many students and teachers
participate in multiple forms or iterates of the designed artifact or emerging practice, making it difficult to know which version of the artifact, or changed practice, caused the observed change or learning on the part of the participants. In their chapter, Sloane, Helding, and Kelly explore the possibilities for between-student modeling over time to work around the lack of a naturally occurring counterfactual in these emerging research designs, with the goal to improve the quality of warranted claims in design research.

Learning occurs in nested data settings: students in classrooms, or learning groups, or classrooms in schools, for example. These “nestings” create problems in conceptualization, design, and analysis. While much design research is qualitative in nature, Sloane (this volume) describes how quantitative researchers have begun to deal with nested data structures and the complexities of building theory and drawing inferences when data have this nested structure. In this chapter, the hierarchical linear model (HLM) is described and is shown to provide a conceptual and statistical mechanism for investigating simultaneously how phenomena at different organizational levels interact with each other.

Two themes in design research include the paths taken and not taken during the design process (Bannan-Ritland & Baek, this volume), and the relationship of parts to the whole intervention process that embrace but extend beyond the idea of nesting described in Sloane (this volume). In his chapter on modeling complexity, Saari (this volume) explores some of the potential applications of fractals, chaos theory, and other formalisms on understanding systems that cannot be accounted for with simple cause-and-effect descriptions. Both this chapter and the one by Hjalmarson and Lesh argue that perturbing a system, and predicting outcomes and measuring them is not neatly linear. Indeed, when the intervention becomes a part of and influences the system, the selection of a good control condition becomes difficult.

Tools to Build a Design Research Community

In addition to efforts to build a community of design researchers, such as exemplified by this book and the one by van den Akker et al. (2006), Kali’s chapter proposes a mechanism (the Design Principles Database) to enable researchers to systematically add their design principles to a database of other lessons learned. The Design Principles Database encourages community building, in which researchers build on each others’ knowledge of design, articulated as principles for design, to create new designs. Researchers explore the application of these principles in new contexts and bring their findings back to the network. In this manner, knowledge about design grows in the community, and design principles are debated, refined, or warranted with additional, field-based evidence.

Technological Affordances

Much is made of the affordances of technologies, with often extravagant claims about how each new technology will revolutionize education. Zaritsky’s chapter discusses how affordances or features (in this case of DVDs) that might impress a developer may not influence learning as expected. Additionally, media grammars (cinematic techniques) were utilized in his project, but often went unnoticed by users. Zaritsky argues that small focus groups of prospective users of technology are highly informative about the likely impact of technologies and should be used more widely. In short, Zaritsky stresses the importance of confirmation of impact of technologies rather than
enthusiastic presumption. In this way, we are reminded of the crucial step of the deployment, or implementation, of the design, as stressed by Ejersbo, Engelhardt, Frølunde, Hanghøj, Magnussen, and Misfeldt (this volume) in their osmotic model of design research.

**Scaling and Diffusion**

Scaling research looks at how innovations move from small to larger implementations, while being actively promoted by some external source. Diffusion research looks at how innovations are adopted, adapted, revised, or rejected as part of a social process among some identified community, often without direct external promotion (Rogers, 2003).

Brazier and Keller (this volume) provide a model to underscore how multiple stakeholders with different objectives support decision-making regarding innovations that are often at odds with and do not support the goals or desires of the innovators. Prioritizing the needs of different stakeholders in the system is often a political process, and can lead to retrograde consequences from the innovator’s point of view.

Wolf and Le Vasan (this volume) argue that schools are not uniform organizations. Some may be fertile ground for adopting an innovative practice; others, for a variety of reasons (e.g., lack of administrative support, the press of academic testing) are not. They show that in doing design research in Singapore, it matters that a school, and its organizational climate, provide conditions for the innovation to survive among teachers. The authors provide an instrument that may be used to gauge the “readiness” of a school to be successfully involved in a design research project.

The chapter by Ejersbo et al. introduces an “osmotic” model to underscore the interplay between designing and reflecting on theory. By the use of three examples, the authors demonstrate (as others do in this book) that actual design research projects are more complex and nuanced than simple descriptions might suggest. Of course, this remark applies equally to any research method, even stylized methods such as randomized field trials (e.g., National Research Council, 2004).

Drawing on a unique record of building toward the scaling of an innovation, Roschelle, Tatar, and Kaput (this volume) point out that neither traditional program evaluation research nor design research are adequate to the task of understanding how to bring an innovation to scale. Part of the problem is that there may be no prior adequate measure of success or that success may be defined differently in different settings.

The authors describe the research trajectory of SimCalc, an innovation that supports learning the mathematics of change and variation, as occurring in six phases, each moving successively closer towards a scaled project. The phases fit either into the context of design or the context of implementation. Most strikingly, they point out that over-specifying an innovation can restrict its adaptability, and consonant with Rogers (2003) they show that an adoptable vision cannot be too radically different from existing practice within the community of potential adopters.

Another unique long-term project that illustrates recursive cycles of development, intervention, and re-design of the intervention was conducted by the MIND Institute for a project called M + M (Math + Music) during the academic years 1998 through 2004 (Martinez, Peterson, Bodner, Coulson, Vuong, Hv, Earl and Shaw, this volume). The keyboard component was designed to teach basic musical concepts and skills necessary for playing the piano. Independently and in groups, students learned a repertoire of progressively more difficult songs. STAR (spatial-temporal animated reasoning) software was designed to develop skill in transforming mental images to enhance
spatial-temporal abilities and understanding of mathematics concepts. The transformations involve symmetry operations applied to two-dimensional figures. Other computer-based games challenged children to apply their spatial-temporal skills to solve mathematics problems in particular, problems involving fractions, proportions, and symmetries. The data show that a spatial-temporal approach to learning key mathematical concepts, allied with music instruction, can produce gains in proficiency with mathematics concepts and skills among children who are academically at-risk. Using a largely non-verbal approach to teaching mathematics concepts, the M + M intervention produced an overall benefit on mathematics achievement for participating second graders in comparison to control group students.

The cumulative findings of this multi-year research project imply that a large segment of students, perhaps most, could benefit from an approach to learning mathematics that appropriates spatial-temporal reasoning along with music training. The use of spatial-temporal reasoning and representations might hold special promise for English language learners because of its relative de-emphasis of language. The M + M project was not a static intervention, but instead evolved through feedback over the course of its implementation. In aggregate, the M + M project demonstrates the viability of the design research approach to educational interventions for advancing students’ learning and the theories on which effective interventions are based.

Conclusions

We have attempted, by use of the device of the “commissive space” and this cross-cutting review of themes, to position existing and emerging discussions on research methodology in a climate of engagement and discussion. While we value critiques of all methods, we eschew criticisms that can only lead to sterile debate and provincialism. A review of the chapters in this book will illuminate the proposition that the scientific study of interventions in complex educational contexts remains in its infancy, but that visible progress is underway. We hope that the reader will find useful the highlighting of cross-cutting themes that differ from the nine-part structure of the book. In particular, we see the importance that assessment plays in guiding iteration in design research cycles. We also described: (a) the role of documentation in illuminating theory, (b) the availability of statistical models that match the messiness of design research projects, (c) the progression of knowledge as a design research field through use of database tools, (d) the testing of affordances as an integral part of design research, and (e) the issues of scaling and diffusion that impact how design research is implemented at higher levels of organization. What remains is the pleasure of reading the individual chapters of the many authors who responded to the general challenge: how may the processes of design and research intersect and cross-pollinate to advance education?

References


