Design-Based Research in Education

A white paper prepared by

William A. Sandoval
University of California, Los Angeles

For the Committee On
Designing Citizen Science to Support Science Learning
The National Academies of Science, Engineering, and Medicine
Board on Science Education

December 1, 2017

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

Introduction

Educational design research, also commonly known in the U.S. as design-based research (DBR), is an approach to studying educational interventions that arose in the early 1990s as a means to solve a problem of research on learning: how to understand the effects of an intervention to promote a form of learning, when the learning processes themselves were not well understood. This problem arose, in part, as psychologists encountered the inherent difficulties in trying to translate findings from laboratory research into instructional interventions in the "blooming, buzzing confusion" (Brown, 1992, p. 141) of classrooms. Such psychologists, as well as artificial intelligence researchers, cognitive scientists, and cognitive anthropologists, were beginning to coalesce into the nascent field of the learning sciences At the same time, Collins (1992) argued that research on educational technologies had to move toward a design science, through a method he called "design experiments." Such experiments would focus, ideally, on tracing the functions of educational technology interventions, not just that they work to achieve an outcome, but how they work. Collins argued this was the only way to understand how seemingly successful designs could be transported across multiple places with some likelihood of maintaining success.

There are many fields that practice their own versions of design research or design studies. Engineering fields engage in research on specific designs not just to improve specific products but to develop processes of engineering, and aeronautics in particular has been used as a helpful analogy for DBR (Collins, 1992; O'Neill, 2012). Given the many different fields engaged in design and the varieties of approaches with which specific designs and the practice of design might be studied, it would be a complex task to examine how to demarcate DBR from other

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

similarly named pursuits. The main aim here is to characterize the broad nature of DBR in

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

education, as it has been practiced and analyzed over the last 25 years, such that researchers of citizen science projects can determine when and how the approach might benefit their work. The bulk of this paper is spent describing DBR in such terms, at quite a bit of detail, to enable citizen science researchers to make sense of the approach for themselves. The concluding section poses some questions citizen science researchers probably need to consider to decide when and how DBR is sensible to use. These are very much questions posed from a DBR insider who is also a citizen science outsider.

DBR has gone by many names. Collins (1992) coined the term "design experiments" to label what he saw as a method of developing the design science of education he envisioned. Brown (1992) popularized the term in a seminal article for the then-nascent field of the learning sciences, and the approach became more widely known in education research during the 1990s (Collins, 1999). As the approach became better known, the term "design experiments" became confused with notions of experimental design. At the turn of the millennium, researchers in the learning sciences began to refer to "design-based research" (Design-Based Research Collective, 2003; Hoadley, 2002) or even just "DBR" (Edelson, 2002). Europe had a long history of "teaching experiments" (Steffe & Thompson, 2000), and a notion of "developmental research" (Gravemeijer, 1994) concerning the development of pedagogy (whereas in the United States, that phrase almost always refers to research on children's cognitive and/or social development). Today, this general mode of research is known either as educational design research or design-based research (DBR). I will use DBR throughout the remainder of this paper for brevity.

A fundamental thing to understand about DBR is that it is not a research methodology, per se. Methodologies have specific methods, and it is not at all clear that DBR can be

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

characterized by a distinctive set of methods, but rather by a set of social and epistemological

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

commitments (Barab, Dodge, Thomas, Jackson, & Tuzun, 2007; Bell, 2004; Bereiter, 2002). In the last two decades there have been quite a number of books (Dai, 2012; Kelly, Lesh, & Baek, 2008; McKenney & Reeves, 2012; Van den Akker, Gravemeijer, McKenney, & Nieveen, 2006) and journal special issues (Barab, 2004; Kelly, 2003; Sandoval & Bell, 2004) attempting to map the space of DBR. These efforts mainly illustrate Bell's (2004) observation that DBR spans wide theoretical and methodological breadths. Nearly all of this writing on DBR as a mode of research comes from practitioners of DBR, as reflections on practice.

DBR as an approach to innovation

The key to understanding DBR as an approach both to intervention and to research is to understand it as fundamentally about sustained innovation (Bereiter, 2002). Looking across a range of writing on DBR, it can be characterized by five social and epistemic commitments (Sandoval, 2013).

Joint goals of improving practice and refining theory. DBR understands learning as movement along trajectories of thinking, trajectories that are sensitive to perturbation and tuning by instruction (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). DBR assumes such trajectories are interactional and contextual. Rather than seeking universal principles of learning that are presumed to hold for all people everywhere, DBR assumes a view of learning settings "not as deterministic, but as complex and conditional" (Confrey, 2006, p. 139). The interest on understanding the interactional relations between ways of thinking and forms of instruction combined with the assumption of conditional localization leads toward design-oriented approaches to the study of learning. As Confrey and others since Brown (1992) have pointed out: the conditions of learning often have to be designed and created in order to study the very learning processes themselves.

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

Sustained engagement with stakeholders. DBR is characterized by sustained engagement with the people designs are intended to serve. The nature of such engagement can vary widely, but DBR typically is carried out in collaboration with an audience rather than done to that audience. Stakeholder collaboration is most commonly seen in the design and enactment phases, with few examples in the DBR literature on collaborative analysis. Collaborative partnership with relevant stakeholders derives from assumptions of contextualization. On one hand, because DBR focuses on innovation, designs are often under-specified and fuller specification happens only during enactment, and necessarily in collaboration with partners who have agreed to work with a research team. In some approaches to DBR, the intervention itself emerges in response to local conditions more than it is developed in advance (Cobb, 2001). In other efforts, the nature of engagement seems more like that of other forms of intervention research, where a design is created by a research team and relevant populations are recruited to use the design and its use is examined with attention to local variations.

Iterated cycles of design, enactment, and analysis. Nearly all discussions of DBR stress its iterative nature. Design itself is inherently iterative, as early prototypes are tested, formally or informally, to see how well they meet specifications, and aspects of designs are repeatedly tinkered with until a design meets some criteria. DBR is typically described as occurring through cycles of design, enactment, and analysis; following an analysis of some enactment, a design is revised and enacted again, and so on. How these cycles are constructed, and how the separate phases are conducted can vary widely. What is being designed and revised, where or how it is being enacted, how it is analyzed, and how each cycle differs from others, can vary widely.

Methods that link processes of enactment to outcomes. A key feature of DBR is its explicit attention to connecting processes of enactment to observed outcomes, both intended and unintended. This is why DBR researchers refer to enactment rather than implementation. Implementation connotes a static view of the design and an unproblematic view that designs can be evaluated simply to see whether or not they work. The core assumption underlying DBR is that the relevant questions are about how a design works in particular settings. To generate causal mechanisms to explain how a design works, and particularly to identify possible underlying principles that may fundamentally characterize a design solution, requires that potential causal processes be observed and connected to outcomes of interest.

DBR typically assumes the processes of enactment observed in one context cannot be presumed to generalize to other contexts. Rather, the notion of context becomes problematized. One issue is to be mindful of the permeable boundaries between the designed learning environment developed by researchers outside of any particular setting, and the existing features of a particular setting that have already been designed to support learning in that place (Tabak, 2004). Tabak warns that a failure to understand the existing (endogenous) design can lead to a fundamental misunderstanding of how the researchers' (exogenous) design functions.

Strives to produce usable knowledge. DBR is centrally concerned with the production of what Lagemann (2002) calls usable knowledge – knowledge about learning and teaching that is practically useful for the people involved in learning and teaching in everyday settings. Bell (2004) discusses the theoretical breadth evident in the spectrum of DBR, and notes this breadth includes variability along a continuum of nomothetic (seeking generalizable laws) to idiographic (characterizing specific cases) efforts. Along this continuum, DBR researchers have claimed that the approach can generate different kinds of knowledge, pertaining to either the process or outcomes of design or the processes of learning a design is intended to support.

Edelson (2002) was perhaps the first to attempt to characterize the different forms of knowledge DBR can produce. He identified three types: domain theories, design frameworks, and design methodologies. The last of these seems to have received very little attention in writing about DBR. A domain theory is an account of how learning occurs in some particular domain. In terms of theoretical improvement being a goal of DBR, this is the kind of theory involved. DBR scholars have, from the start, grappled with the generality of such theories, with various arguments for the "local" nature of domain theories (Brown, 1994; Cobb et al., 2003; diSessa, 1991). The sources of localization can vary widely, depending primarily on researchers' theoretical orientation. DBR coming from a cognitive science or developmental psychology perspective (see Bell, 2004) may attribute localization primarily to the difficulty of disentangling observed learning processes from the conditions under which they are observed, leading to exhortations to more fully characterize multiple levels of interacting variables or factors to identify generalizable features (Collins, 1999). On the other hand, DBR from cultural psychology or anthropological perspectives sees localization as fundamentally inherent to human activity. Consequently, the research effort is not to adduce general claims out of local variations, but to explain local situations in their own terms (Engeström, 2008). Such efforts often focus less on measuring outcomes from an intervention, and more on how designs are interpreted by participants, how they are taken up or resisted in particular settings (Bell, 2004).

A design framework is "a coherent set of design guidelines for a particular class of design challenge" (Edelson, 2002, p. 114). Design frameworks specify particular kinds of design

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

solutions. Any particular design framework is comprised of design principles, and these principles can be developed prior to or as an outcome of DBR. Recent efforts have been made to articulate design principles in a more general way, to promote the spread of design knowledge across research projects (Kali, 2006; Kali & Linn, 2008). The idea is that design principles can, at least sometimes, be used independently of a larger framework.

DBR advocates emphasize that these forms of knowledge are usable because they are grounded directly in efforts to design for specific forms of educational practice. This argument plays out on two levels. At the local level of any particular setting, systematic research-driven efforts to change practice are likely to generate usable insights about that setting for the participants in the work. At a higher level, because DBR is driven by problems of practice and occurs in settings of practice, its insights are more directly available to be applied to practical settings. The strength of DBR is often identified as theoretical innovation grounded in specific experiences to create change (Brown, 1992; Cobb, 2001; Edelson, 2002; Steffe & Thompson, 2000).

General methodology of DBR

Given the variety of work described as DBR, it would be misleading to argue that there is a single way to do it correctly, or even a finite set of ways to do it. It may not even be desirable to articulate a single approach to the conduct of DBR (Bell, 2004). On the other hand, there has been concern both inside (Dede, 2004; Kelly, 2004) and outside (Levin & O'Donnell, 1999; Phillips & Dolle, 2006; Shavelson, Phillips, Towne, & Feuer, 2003) the DBR community about the standards of DBR and the strength of claims that might be possible through such work. From the writing on DBR over the last two decades, some features of a general methodology can be discerned and articulated as they pertain to both the design and the research aspects of DBR.

Problem analysis

DBR needs a design to study, and a design needs an analysis of the problem to be solved. Consequently, DBR begins with an analysis of the learning problem to be solved (Edelson, 2002). To analyze a problem, or an opportunity, requires that it be defined. Educational problems are typically defined in terms of some need experienced by some group. Educational needs, of course, are expressed in a variety of ways by a wide range of stakeholders. Standards documents express aims for segments of educational systems, and consequent needs. Prior research can describe the needs of various sub-groups within education. A needs assessment can also be done through direct observation or interaction with particular groups who express their own need.

Whatever the origin of the problem or opportunity, good problem analysis includes a review of research literature relevant to the problem. Relevant research helps DBR teams to characterize the problem they are trying to solve, it can help to formulate goals, or even generate design ideas. A good problem analysis generates design ideas that go beyond current practice to focus on potential. As Bereiter (2002) puts it, good DBR has a "visionary quality" focused on innovation that goes well beyond notions of program improvement or increasing efficiency – an ethos of potential and what might be rather than what is or what has been done.

Initial design

A problem analysis leads to an initial design. The form of this design can vary widely and depends upon a number of factors. Crucially, DBR sees design as an inherently theoretical activity, as a means of embodying theoretical conjectures about learning in the chosen problem context and how such learning can be supported (Cobb et al., 2003; Confrey & Lachance, 2000; Sandoval, 2004). General descriptions of the pursuit of this embodiment can be hard to come by,

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

as the acknowledgment that there is no single best way to go about design leads many writers to describe DBR through their own examples (Barab, 2006; Confrey & Lachance, 2000; Middleton,

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

Gorard, Taylor, & Bannan-Ritland, 2008). These exemplars are presented as models to illustrate how particular theoretical commitments to learning are inscribed in specific design choices. Within recent writing about DBR, however, one can find a few efforts at a general description of the method of design.

Cobb and his colleagues have described their own process of hypothesizing learning trajectories as their approach to design (Cobb, 2001; Cobb et al., 2003; Gravemeijer & Cobb, 2006). They start with a problem analysis that strives to identify the starting points of learners and specifies the desired endpoints of instruction, and then conduct what they describe as thought experiments to conjecture trajectories of learning or development. Yet, even in their most recent reflection on their method (Gravemeijer & Cobb, 2006), they do not describe in detail the character of such thought experiments. The main precept to be drawn from this approach is the articulation of presumed starting points and desired endpoints as metrics for judging progress. Particularly as this approach to design is responsively emergent to issues that arise during enactment (Cobb, 2001), it is crucial to have clear endpoints in mind.

Middleton and colleagues (Middleton et al., 2008) have described an approach to design drawing directly from engineering, and identify the job of design as describing relationships between the form of design elements, the behavior of those elements, and the functions such behaviors should produce within the designed system. Although they exemplify their process through describing a particular project, their description of the overall process is somewhat abstract. Yet, as it applies particularly to the development of educational technologies or other material tools or resources, their description quite usefully highlights the imperative to specify how specific features are conjectured to lead to particular behaviors and functions.

Sandoval (2004, 2014) has articulated an approach to mapping the conjectures derived from a problem analysis through a design. This approach distinguishes four kinds of design elements: material tools and resources, including technologies or other curricular materials; task structures, specifications of the goals and demands of the tasks through which designed tools will be used; participant structures, specifications of the roles and responsibilities that participants take on during tasks; and discourse practices, ways of talking that can be promoted to productively influence how task and participant structures are taken up. Conjecture maps then articulate how these design elements are predicted to interact in their intended setting to produce mediating processes: forms of interaction that are evidence that a design functions as intended. Designers can further specify how predicted mediating processes lead to desired outcomes. Conjecture maps thus distinguish between design conjectures, hypotheses about how a design will work to produce mediating processes, and theoretical conjectures, hypotheses about how mediating processes lead to intended learning outcomes.

It is quite evident from writing on DBR since Edelson's (2002) effort to describe a process, that DBR scholars are simultaneously aware of the desirability of models for the process and the difficulty of articulating such models in ways that acknowledge the variability in good designs and good research. The character of all of these approaches that distinguishes DBR from more traditional instructional design work is the view that designs are not merely informed by theory (cf., Carroll, 1993), but embody theoretical commitments that are targets of empirical scrutiny.

Enactment

A general description of the enactment phase of any DBR project or study is extremely hard to produce, because the nature of designed interventions varies so widely, and the nature of

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

enactment differs greatly depending upon where a DBR team might be in the trajectory of their work (see below). Some of the general methodological issues specific to DBR are addressed below, but with respect to the enactment phase of DBR it may be worth pointing out that both the enactment of the designed intervention and the enactment of research plans can be dynamic, and subject to change as they progress. A common, and perhaps unsettling, feature of many DBR projects is that both the interventions designed and the research on their function can be emergent, especially during the early stages of a DBR program.

Cobb and Gravemeijer have articulated a highly emergent approach to DBR, where the designed intervention is in many ways constructed during enactment through what they describe as micro-cycles of design and analysis (Gravemeijer & Cobb, 2006). Their approach can include the development of tools, such as software, that are not open to change during enactment, but in general Cobb and his colleagues' approach to design is purposefully emergent: as learning trajectories are hypothesized, instructional trajectories are hypothesized, but these instructional trajectories remain very open to change in response to what happens daily in settings of enactment. What is learned during one day of enactment is used to modify what happens the next. This produces detailed traces of how interventions unfold over time (Cobb, 2001). This is a particular form of DBR, it must be noted, and not characteristic of all DBR work. Many DBR projects produce designs that are more fixed for particular enactments, and are most easily revised between enactments (e.g., Linn & Hsi, 2000).

The emergent nature of designs in DBR necessitates flexibility in research methods and data collection. Early design experiments tended to follow an ethos of collecting as much data as possible in order not to miss anything that might turn out to be important (Brown, 1992). Besides

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

being impractical, this too easily leads to unsystematic data collection and analysis, and an

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

"anything goes" feeling that undermines the credibility of research (Kelly, 2004). It also obscures the problem that any form of data collection enacts theoretical perspectives, even videotaping (Hall, 2000), which Brown singled out as a way to capture everything going on in a classroom. Conjecture mapping developed as a means to articulate commitments to processes and outcomes and ways to measure or observe them (Sandoval, 2014).

Analysis

The general cycle of DBR frames analysis as something that happens after enactment, although the preceding section makes clear that ad-hoc analysis is a common feature of DBR enactments. This phase is a retrospective analysis (Gravemeijer & Cobb, 2006) of the data collected during an enactment for the purpose of thoroughly understanding how a design functions. This analysis involves testing both design conjectures and theoretical conjectures (Sandoval, 2014).

Design conjectures are claims that elements of a design will function to produce certain interactional processes during enactment. That is, if learners use features of a design in intended ways, then they will interact – do and say things – in specific ways. If intended interactions are not observed, then the question is whether or not learners used design features as intended. If not, then the design probably needs to be revised: task or participant structures made clearer, or tools or materials improved, and so forth. If learners use a design as intended but this does not produce desired forms of interaction, then this suggests the underlying design itself is flawed and has to be rethought.

Theoretical conjectures pertain to the underlying theory driving the design. They are of the form: if the design functions as intended (the design conjectures are right and desired interactions are observed), then targeted learning outcomes will occur. Here, then, if the

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

predicted forms of interaction do occur but the intended outcomes do not, the underlying theory is either wrong or underspecified (see Sandoval, 2014).

Methodological challenges for DBR

DBR shares the challenges of any empirical research, including sampling, measurement, generalization, and the various biases that can influence research. Some of these challenges present themselves in particular ways within DBR. The specific methodological challenges discussed here are facets of the overarching challenge for DBR: accounting for the function of a design for learning as enacted in context. Defining and accounting for contextual variations and their influences on the conclusions that can be drawn from DBR have been the focus of much of the writing about the approach (Barab, Hay, & Yamagata-Lynch, 2001; Brown, 1994; Cobb et al., 2003; Collins, 1999; Collins, Joseph, & Bielaczyc, 2004; Hoadley, 2002; Steffe & Thompson, 2000; Tabak, 2004). The discussion here addresses this overarching concern in terms of some concrete issues of method that are not often discussed, amplifying (Sandoval, 2013).

Site selection

Sampling is an issue in any research design, and how samples are chosen has a serious influence on the validity of claims made from the research. In the experimental tradition of hypothesis testing to which DBR is often contrasted (Hoadley, 2004; Kelly, 2004; Shavelson et al., 2003), sampling is primarily concerned with establishing representativeness with respect to an underlying population so that any observed effects can be generalized. In randomized experiments or quasi-experiments without random assignment, the question of interest is usually framed as "what works?" (Raudenbush, 2005). For that question, sample selection is key to establishing the equivalence of treatment and comparison groups in terms of both their background characteristics and their potential to benefit from the experimental treatment.

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

Satisfying these conditions increases researchers' confidence in the inferences produced by statistical tests of effects. These concerns can be a factor in DBR as well, but the acute issue of sampling for DBR is selecting sites for research where the problem one is trying to solve exists, or the opportunities one is trying to exploit can be realized.

DBR requires sites friendly to the research effort. It is usually crucial to choose sites that are open to innovation and have the interest and capacity to collaborate at the level required by the research team. Site selection in the early stages of DBR is therefore more about selecting collaborators than sites. Moreover, collaborating practitioners can be more than simply willing participants, they sometimes can significantly contribute to the development and sustainability of innovation (Bereiter, 2002). Even though the nature of collaborations varies between DBR efforts and changes over the trajectory of a research program, willing collaborators remain crucial to success. This is true whether the collaborator is a single classroom teacher willing to work with a research team over several years (e.g., the CLP/KIE project, Linn & Hsi, 2000) or an entire school district (Fishman, Marx, Blumenfeld, & Krajcik, 2004). The collaborative nature of DBR highlights that site selection affects not just issues of sampling, but influences the designed environment directly because of the inevitably blurry intermingling of the external design and the existing, ongoing designed and emergent features of the setting into which designs are imported (Tabak, 2004).

Instrumentation and outcome measurement

The core aim of DBR to link targeted outcomes to the functional processes created through designs raises perhaps the central methodological challenge of DBR. The variability in approaches to DBR, and the wide spread in researchers applying the label to their own work has

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

led to a concern that DBR lacks an argumentative grammar to describe how such links can be warranted (Kelly, 2004). The theoretical variety among approaches to DBR (Bell, 2004) probably makes a singular description of such a grammar unlikely. Part of this variety is that processes and outcomes can be observed and measured in many different ways, and choosing the best techniques of measurement or observation depend upon the questions driving the research.

Bereiter (2002) points out that the best DBR has a visionary quality that imagines possibilities that do not yet exist in current practice. These possibilities can be roughly broken down into two kinds: new pedagogical possibilities for achieving currently desired outcomes, or new outcome possibilities and the pedagogies that might produce them. Both kinds of visions have been pursued in DBR, and each presents measurement challenges to solve.

A good deal of DBR aims to transform how people achieve well-established outcomes. In such efforts, the primary measurement challenge is the selection of appropriate outcome measures from available alternatives. The challenges in this situation are to choose one or more established measures of the outcome of interest and adapt them as necessary. Such efforts are relatively common in mathematics and science domains where concept inventories and the like have been extensively developed. An advantage of choosing such off-the-shelf measures is they allow for more direct comparison of an intervention with relevant related work, and they are often clear to important stakeholders.

It is just as likely, however, that DBR teams are interested in transforming both how students learn and what they learn (Roschelle, Pea, Hoadley, Gordin, & Means, 2000). This creates the more difficult challenge of designing instruments that can measure the outcomes of interest. Schwartz, Chang, and Martin (2008) argued that DBR can profit from more attention to issues of instrumentation and measurement. They distinguish between apparatuses to make effects happen and measures of those effects, drawing from science studies of the role of

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

instrumentation in the advancement of science. They argue DBR has focused primarily on the development of apparatuses to perturb environments to engineer forms of learning (Cobb et al., 2003), but tend to neglect the design of measures of the outcomes of such perturbations. The design of good measures, they argue, is crucial for generating evidence appropriate to the innovations of DBR, for the simple reason that more widely available measures are tuned to current practice rather than possibility. Schwartz and colleagues describe a process of "assessment experiments" they use to tune measures to innovations.

Characterizing relevant processes

Assuming a DBR team has access to good outcome measures, the enterprise requires methods to characterize the processes at work in a learning environment that are causally responsible for those outcomes. These processes are sometimes described from a design perspective in terms of functions of the design (Middleton et al., 2008), or as interactional processes that mediate production of desired outcomes (Sandoval, 2014). The observation of mediating (functional) processes is an indication that a design functions as intended, whereas the observation of the ultimate outcomes targeted by an intervention indicates whether or not that functional design achieves its intended aims (Middleton et al., 2008 refer to these functional processes as "behaviors," and intervention outcomes as "functions").

There are basically two ways mediating processes can be observed during the enactment of a design: they can be observed directly, or they can be observed through proxy measures of interaction (Sandoval, 2014).

Direct observation. There are a wide range of methods of observation developed in the social sciences, broadly, and taken up in educational research. Particular approaches in DBR are explicated in recent texts (Kelly et al., 2008; McKenney & Reeves, 2012). Video data, however,

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

has been a common fixture in DBR since Brown's (1992) seminal discussion of her approach. There are several features of video recordings that lead DBR teams to prefer them as data sources for capturing and analyzing relevant processes of design function. One is that video records capture a complete veridical record of the activity within the frame, generally with much more detail than a human observer can record through other means. Second, this detailed record is captured over some temporal span that can be viewed repeatedly. This allows potentially important interactions to be scrutinized in great detail, their emergence to be traced within the setting, and their effects traced forward in time. Third, analysis can be done by multiple analysts, reducing potential observer biases and supporting triangulation. These advantages come at some cost, as collecting, storing and analyzing video is expensive relative to other methods. The centrality of video data to DBR has produced a comprehensive guide in the learning sciences community (Derry et al., 2010).

A range of designs using software tools enable the collection of detailed logs of user interactions that can be useful for tracing how those tools are used in enacted learning environments. These sorts of logs are easy to create, usually, but, like video, they produce huge amounts of data. An entire field of learning analytics has sprung up over the last decade to develop methods for mining such data for insights into how design functions (Larusson & White, 2014).

Proxy observation. A second means for observing how any particular design functions in a particular setting is through the analysis of artifacts produced during the enactment. The BGuILE project, for example, was designed to support the development of scientific explanations as a means to learning about natural selection and about the nature of science

(Reiser et al., 2001). Analyses of the explanations students wrote during instruction showed

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

specific ways that particular design versions did and did not support students. For example, making data citation an explicit task demand encouraged students to cite data, but did not help them clearly relate data to claims (Sandoval & Millwood, 2005).

Critics of DBR worry that designers are biased toward showing their intervention works (Shavelson et al., 2003). On the contrary, DBR demands attention to negative results or the aim of sustainable innovation cannot be reached (Bereiter, 2002). Design is iterative by nature because early designs rarely solve their target problem fully, and the impetus of design is improvement. This naturally tunes DBR to shortcomings in designs and focuses analytic attention on their causes and possible solution (Sandoval, 2004). This is how DBR advances theoretical insight, by driving analyses of processes of design enactments that refine the conjectures embodied in designs.

Working across levels of analysis

DBR takes place in authentic settings where learning occurs; that is, pretty much everywhere. There is, of course, tremendous focus on learning in schools, but DBR occurs in and about museums, after-school clubs, and workplaces. Even DBR focused on schooling occurs in myriad settings, such as classrooms, professional development meetings, and local education agencies. Obviously, each of these settings is comprised of individuals who participate in them, and their very participation is itself embedded within social, institutional, cultural, and historical contexts. It is now taken for granted that all of these layers influence how any educational intervention might function (Collins, 1999). Understanding how a design functions in a particular setting, and how that function could vary across settings, requires attention to how the researchers' level of analysis interacts with adjacent levels of analysis.

Collins (1999; Collins et al., 2004) has articulated several levels at which particular designs could be analyzed, moving from the individual cognitive level through to the institutional level (e.g., the school), and lists a range of variables DBR might attempt to measure, including climate and learning variables that could characterize individuals or groups, and system level variables at the institutional level. It is not at all clear, however, that any single DBR effort within a particular research cycle could attend to all of these levels at once. Collins et al. (2004) suggest that multi-disciplinary teams with expertise distributed across all of these levels could do the job, but it may not be feasible, especially in the early stages of a DBR effort, to engage such teams. Bell (2004) points out that the theoretical breadth observed across the landscape of DBR includes a wide variety of aims and it is quite possible that different aims would appropriately lead researchers to focus on one level of the educational system over others.

At the same time, as DBR is interested in the sustainability and diffusion of innovations, there would seem to be a minimal set of levels of analysis for most, if not all, DBR. These levels are articulated by Rogoff (1995) as inter-related planes of analysis: the personal, the interpersonal, and the community. The personal plane focuses on how individuals experience an educational intervention. The interpersonal plane focuses on the interactions between individuals, and the community plane on the norms and practices of the setting that affect interactions. Learning environments could be designed to directly affect any and all of these planes, but even if they were focused only on one plane research demands some characterization of adjacent planes to adequately characterize the function of designed elements in the learning environment.

Coordinating across planes of analysis requires methods tuned to each plane. Collins et al

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

(2004) make very specific suggestions of the kinds of data and analyses best suited to characterizing different levels, and these may be particularly useful for newcomers to DBR, although it may well be that these or any specific set of prescriptions would be contested from other theoretical perspectives. A methodological issue that can be easily forgotten is that these different planes of analysis operate at different time scales (Lemke, 2000), and research methods must be tuned to these differences. For instance, community level norms are relatively stable and usually tacit, so brief records of interpersonal interaction may not surface them. Rather, observing the function of norms within a community and their influence on interpersonal interactions requires longer scales of observation (e.g., Cobb, Stephan, McClain, & Gravemeijer, 2001).

Causal attribution

While a great deal of educational research is concerned with the demonstration of causal effects, DBR aims to explicate causal processes (Sandoval, 2014). This comes from a view of causality that Maxwell (2004) describes as scientific realism, and argues that causes can be observed in nature, and not merely inferred from the regular co-occurrence of events. This notion of causality is fundamentally multivariate and multirelational. DBR tends not to conceive of elements of a design as a set of variables and interactions, however. Bereiter (2002) argues that analyzing variables and their interactions does not help designers very much: while designs must be decomposed to some extent, the classical researcher tears the design into finer and finer grained details that move further and further away from the design itself. Bereiter and Sandoval rely on Cronbach's (1975) identification of the limitations of variable-oriented approaches to the study of interventions: there are too many variables and way too many interactions. More fundamentally, the elements that make up a design are not easily captured as variables, as elements are design to function together. Some software tool, for example, is not merely a variable, as it is designed to be used in a certain way, that puts people in particular interactional configurations, and so on.

The emergent grammar of design research focuses on clear articulations of the design and theoretical conjectures at play in a designed environment, and on methods that can test them (Sandoval, 2014), as suggested above. Descriptions of participants' interactions with designs is necessary to enable researchers to trace the "crucial components and relations" of specific designs (Engeström, 2008, p. 4). This happens by attempting to characterize functions rather than effects, functions that arise from complex interactions among elements of a designed environment and the people who interact with them (Sandoval, 2014).

Trajectories of DBR

DBR cannot be understood as a method for conducting any single research study. Everyone writing about DBR as an approach or methodology emphasizes its iterative nature. Yet, not much of this writing has attempted to describe the trajectories that iterating cycles of DBR might play out, or the kinds of questions different trajectories might enable researchers to answer. Given the highly dynamic nature of design, the variety of theoretical perspectives on learning evident in DBR, and the difficulties inherent in attempting to create and observe causal processes of learning emergent from design variants, trajectories of DBR can unfold in quite different ways. A survey of DBR projects and writing about DBR suggests several possible trajectories of DBR programs.

Scale

A number of early DBR efforts, especially those focused on educational technologies and often grounded in the field of instructional systems design, followed trajectories of increasing scale. This is a trajectory common in the broader work on educational interventions. The initial

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

focus of work is on the design of an intervention and pilots of that to refine the design itself. These pilots may be conducted in the intended setting of use, or simplified proxies of that setting. Once a stable design is achieved, it is taken into its intended setting of use and studies aim to show it achieves its intended effects through hypothesized functions. This leads to larger scale efficacy studies that, if successful, enable the dissemination of what, by now, is considered a finished product. A concise summary of this trajectory is provided by (Bannan-Ritland, 2003) and a rather detailed treatment of work along this sort of trajectory can be found in (McKenney & Reeves, 2012).

This is a trajectory of design and research that locates DBR specifically at the start of the trajectory and renders it inapplicable as the scale of study grows larger and the designed intervention becomes calcified.

Depth

An alternative view of DBR is as a method that allows researchers to explore learning processes in-depth. In this sort of work, designs aim to produce some novel form of learning, and how those forms emerge and develop is studied in detail by introducing variants on the instructional designs intended to promote and support them. Cobb and colleagues' work on the development of sociomathematical norms is an example of this kind of DBR (Cobb, 2001; Cobb et al., 2003; Cobb et al., 2001). Another well-known classroom example is the long running program of Lehrer and Schauble focused on children's development of scientific modeling practices (Lehrer & Schauble, 2000, 2004, 2002; Lehrer, Schauble, & Lucas, 2008). These two efforts share a focus on the co-development of learning environments with practitioners (classroom teachers, in both cases), and are focused more on the design of activity structures and

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

discursive practices than on the design of materials and tools. This trajectory frames design as a

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

means to generate novel forms of learning, even leading to previously unseen forms of cognition or interaction (diSessa & Cobb, 2004; Sandoval, 2004) that require study.

A different form of in-depth DBR is the refinement over an extended period of time of a designed environment in a stable, usually small-scale, setting. A classic example of this is Linn's programmatic develop of the KIE/WISE science learning environments developed over more than a decade through collaboration primarily with a single teacher (Linn & Hsi, 2000). Each design iteration expanded the learning environment in some way, sometimes comparing different forms of tools or materials, to build a comprehensive set of tools, materials, and activity structures.

Context

A third form of trajectory involves studying learning environment designs across multiple contexts, not necessarily in order to take them to scale, but to understand what Cronbach (1975) called "local departures" from an intervention. This path seems rarely taken in DBR, if at all. Certainly, programs of DBR routinely change sites across study cycles. Changes in site selection, however, do not seem to be made in order to allow the study of that change, per se.

DBR on citizen science

A cursory look at citizen science projects reveals them to be highly variable in structure, means of participation, scale, and goal (Bonney, Shirk, & Phillips, 2015; McKinley et al., 2017). In considering how DBR might be applied to research on citizen science projects, there are a number of questions that citizen science researchers seem best poised to answer. Perhaps some have already been answered; hopefully, framing them can aid citizen science researchers in their work. There seem to be at least two broad issues to consider. One, are the goals of a particular

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.

citizen science project consonant with goals of DBR? If the first question is answered affirmatively, then where are the key sites of interaction for any citizen science project design?

Goal alignment

DBR pursues joint goals of improving practice and learning (or instructional) theory. Citizen science holds distinct but related learning goals: participants have to learn how to do the science relevant to the project; it is generally hoped they learn some substantive science along the way (e.g., about bird migration); and there are also hopes that participants learn something about how science actually gets done. If the question of interest is whether such outcomes are achieved, evaluation research methods would be appropriate. Bonney et al. (2015) summarize a range of citizen science efforts that have demonstrated that participants learn along all of these outcomes.

DBR would be appropriate for projects seeking to understand *how* their design produces learning, and *how* that relates to producing such learning more broadly. Embarking on DBR for any citizen science project would entail specifying how elements of the project design are hypothesized to produce particular forms of interaction (design conjectures), and how those interactions are hypothesized to produce specific learning outcomes (theoretical conjectures). A compelling counter example comes from the GLOBE (Global Learning through Observation to Benefit the Environment), which aimed to get school students to collect usable environmental data as a means of learning science inquiry. Evaluations of GLOBE showed that the data collection protocols provided by scientists were highly amenable to being appropriated into typical didactic instruction, rather than legitimate inquiry (Penuel & Means, 2004). A DBR approach to this work would have focused early attention to how protocols and task structures could be designed to promote particular forms of interaction (mediating processes of design function), and how such interaction was conjectured to produce desirable outcomes.

Sites of interaction

Given the explicit goal in DBR of trying to link processes of design function to the accomplishment of outcomes, intended and unintended, it is crucial to the approach to determine how interactions with specific designs can be observed or measured. An obviously important question to ask is, can participants be observed directly? Cardiel, Pattison, Benne, and Johnson (2016) describe their effort to analyze the design of mobile science exhibits deployed at transit shelters in a city in the Northwest. Museum researchers used observation protocols, field notes, and interviews to document and understand the features of exhibit designs that prompted people to interact, or not interact, with exhibits. One thing Cardiel and colleagues learned is that ancillary features of the exhibits captured initial attention and influenced whether people choose to engage or not. They were able not only to improve exhibit designs through two cycles of design, deployment, and analysis, but they were also able to test a model of stages of attention to exhibits. Citizen science projects may be able to engage in some forms of observations of interaction by following the sort of quasi-naturalistic methods employed in this study.

A good deal of citizen science work now seems to rely on online sites, mobile technologies, and crowdsourcing (Newman et al., 2012). These obviously pose many challenges for design, and many opportunities to employ DBR. Participants' use of web-based tools may best be understood through trace logs or other means of capturing interaction. How mobile applications are used and interpreted could be observed through naturalistic studies where some sample of participants were observed interacting with tools, and perhaps asked to reflect on their use. From a DBR perspective, it would be crucial for project teams to first specify how they thought tools and the tasks for which they are designed are hypothesized to promote learning.

Conclusion

The primary aim here has been to describe DBR as an approach to educational design and research in enough detail that citizen science researchers can judge whether, and how, to employ the approach in their own work. Probably the most crucial point to take away from here about DBR is that it is defined more by its goals than by any particular set of methods (Bereiter, 2002). Perhaps the most crucial of these goals from a methodological perspective is the explicit aim to link the processes of design function that mediate how learners interact with a designed learning environment to what and how they consequently accomplish intended outcomes (Sandoval, 2014). Citizen science is much too large and varied a field of work (Bonney et al., 2015) for an outsider to offer prescriptions for the application of DBR. Instead, citizen science researchers have to ask whether the goals of their work are consonant with the five aims of DBR as outlined here, and whether they have the means to engage in the empirical documentation of interactions with designs in use. DBR has been taken up in other informal science spaces, and it is likely that citizen science research can, and perhaps already has, take up the approach as well.

One final thought: to the extent that citizen science research explores public engagement and action as a learning goal, DBR approaches may help to bridge a gap in educational understanding and design. Very little is actually known about how learning the disciplines, including the sciences, in school translates to engagement with and use of those disciplines outside of school (Sandoval, 2016). School science currently tends to be a place where young citizens do school, rather than really do science. Stromholt and Bell (2017) provide one example of how school instruction can be explicitly linked to out-of-school experiences, in their case carefully coordinated and expansive field trip experiences, to promote youth engagement with science in their everyday lives by linking school to life out of school. Citizen science research seems to have the potential to expand such connections even more, not necessarily by coordinating with school instruction but by providing detailed examples of how designs for public engagement and everyday learning produce learning and action outcomes. DBR approaches may facilitate such work in ways that lead to design principles that could be fed back into designing school learning environments with a more outwardly civic orientation. Schools could become one site where young children commonly do science, changing the possibilities for the kinds of citizen science and public engagement with science they may be able to do as adults.

References

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32*(1), 21-24.

Barab, S. A. (2004). Design-based research: Clarifying the terms [special issue]. *Journal of the Learning Sciences, 13*(1).

Barab, S. A. (2006). Design-based research: A methodological toolkit for the learning sciences. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (pp. 153-172). New York: Cambridge University Press.

Barab, S. A., Dodge, T., Thomas, M. K., Jackson, C., & Tuzun, H. (2007). Our designs and the social agendas they carry. *Journal of the Learning Sciences*, *16*(2), 263-305.

Barab, S. A., Hay, K. E., & Yamagata-Lynch, L. C. (2001). Constructing networks of action-relevant episodes: An in situ research methodology. *Journal of the Learning Sciences, 10*(1&2), 63-112.

Bell, P. (2004). On the theoretical breadth of design-based research in education. *Educational Psychologist, 39*(4), 243-253.

Bereiter, C. (2002). Design research for sustained innovation. *Cognitive Studies, Bulletin of the Japanese Cognitive Science Society, 9*(3), 321-327.

Bonney, R., Shirk, J. L., & Phillips, T. B. (2015). Citizen science. In R. Gunstone (Ed.), *Encyclopedia of science education* (pp. 152-154). Dordrecht: Springer.

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

Brown, A. L. (1994). The advancement of learning. *Educational Researcher, 23*(8), 4-12.

Cardiel, C. L. B., Pattison, S. A., Benne, M., & Johnson, M. (2016). Science on the move: A design-based research study of informal stem learning in public spaces. *Visitor Studies, 19*(1), 39-59. doi:10.1080/10645578.2016.1144027

Carroll, J. M. (1993). Creating a design science of human-computer interaction. *Interacting with Computers, 5*(1), 3-12.

Cobb, P. (2001). Supporting the improvement of learning and teaching in social and institutional context. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: 25 years of progress* (pp. 455-). Mahwah, NJ: Lawrence Erlbaum.

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

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

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15-22): Springer-Verlag.

Collins, A. (1999). The changing infrastructure of education research. In E. C. Lagemann & L. S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp. 289-298). San Francisco: Jossey-Bass.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences, 13*(1), 15-42.

Confrey, J. (2006). The evolution of design studies as methodology. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (pp. 135-151). New York: Cambridge University Press.

Confrey, J., & Lachance, A. (2000). Transformative teaching experiments through conjecture-driven research design. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 231-265). Mahwah, NJ: Lawrence Erlbaum Assoc.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30*(2), 116-127.

Dai, D. Y. (Ed.) (2012). *Design research on learning and thinking in educational settings*. London: Routledge.

Dede, C. (2004). If design-based research is the answer, what is the question? *Journal of the Learning Sciences, 13*(1), 105-114.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., . . . Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences, 19*(1), 3-53. doi:10.1080/10508400903452884

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

diSessa, A. A. (1991). Local sciences: Viewing the design of human-computer systems as cognitive science. In J. M. Carroll (Ed.), *Designing interaction: Psychology at the human-computer interface* (pp. 162-202). Cambridge: Cambridge University Press.

diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences, 13*(1), 77-103.

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

Engeström, Y. (2008). From design experiments to formative interventions. In P. Kirschner, F. Prins, V. Jonker, & G. Kanselaar (Eds.), *International perspectives in the learning*

sciences: Cre8ing a learning world. Proceedings of the eighth international conference

for the learning sciences – icls 2008 (pp. 3-24): International Society of Learning Sciences.

Fishman, B. J., Marx, R. W., Blumenfeld, P. C., & Krajcik, J. S. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences, 13*(1), 43-76.

Gravemeijer, K. (1994). *Developing realistic mathematics education*. Utrecht: Cdβ Press.

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

Hall, R. (2000). Video recording as theory. In R. A. Lesh & A. E. Kelly (Eds.), *Handbook of research design in mathematics and science education* (pp. 647-664). Mahwah, NJ: Lawrence Erlbaum.

Hoadley, C. (2002). Creating context: Design-based research in creating and understanding cscl. In G. Stahl (Ed.), *Proceedings of the international conference of computer support for collaborative learning (cscl) 2002* (pp. 453-462). Hillsdale, NJ: Erlbaum.

Hoadley, C. (2004). Methodological alignment in design-based research. *Educational Psychologist, 39*(4), 203-212.

Kali, Y. (2006). Collaborative knowledge building using the design principles database. *International Journal of Computer-Supported Collaborative Learning, 1*(2), 187-201.

Kali, Y., & Linn, M. C. (2008). Designing effective visualizations for elementary school science. *Elementary School Journal, 109*(2), 181-198.

Kelly, A. E. (2003). Theme issue: The role of design in educational research *Educational Researcher* (Vol. 32).

Kelly, A. E. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences, 13*(1), 115-128.

Kelly, A. E., Lesh, R. A., & Baek, J. Y. (Eds.). (2008). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*. New York: Routledge.

Lagemann, E. C. (2002). *An elusive science: The troubling history of education research*. Chicago: University of Chicago Press.

Larusson, J. A., & White, B. (Eds.). (2014). *Learning analytics: From research to practice*. New York: Springer.

Lehrer, R., & Schauble, L. (2000). Developing model-based reasoning in mathematics and science. *Journal of Applied Developmental Psychology, 21*(1), 39-48.

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal, 41*(3), 635-679.

Lehrer, R., & Schauble, L. (Eds.). (2002). *Investigating real data in the classroom: Expanding children's understanding of math and science*. New York: Teachers College Press.

Lehrer, R., Schauble, L., & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development, 23*(4), 512-529.

Lemke, J. L. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity, 7*(4), 273-290.

Levin, J. R., & O'Donnell, A. M. (1999). What to do about educational research's credibility gaps? *Issues in Education, 5*(2), 177-229.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers*. Mahwah, NJ: Lawrence Erlbaum, Assoc.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11.
McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research*. New York: Routledge.
McKinley, D. C., Miller-Rushing, A. J., Ballard, H. L., Bonney, R., Brown, H., Cook-Patton, S. C., . . . Soukup, M. A. (2017). Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation, 208*, 15-28.
Middleton, J., Gorard, S., Taylor, C., & Bannan-Ritland, B. (2008). The "compleat" design experiment: From soup to nuts. In A. E. Kelly, R. A. Lesh, & J. Y. Baek (Eds.), *Handbook of design research methods in education* (pp. 21-46). New York: Routledge.
Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., & Crowston, K. (2012). The future of citizen science: Emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment, 10*(6), 298-304. doi:10.1890/110294
O'Neill, D. K. (2012). Designs that fly: What the history of aeronautics tells us about the future of design-based research in education. *International Journal of Research & Method in Education, 35*(2), 119-140.
Penuel, W. R., & Means, B. (2004). Implementation variation and fidelity in an inquiry science program: Analysis of globe data reporting patterns. *Journal of Research in Science Teaching, 41*(3), 294-315. doi:10.1002/tea.20002
Phillips, D. C., & Dolle, J. R. (2006). From plato to brown and beyond: Theory, practice, and the promise of design experiments. In L. Verschaffel, F. Dochy, M. Boekaerts, & S. Vosniadou (Eds.), *Instructional psychology: Past, present and future trends: Sixteen essays in honour of erik decorte* (pp. 277-293). Amsterdam: Elsevier.
Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher, 34*, 25-31.
Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). Bguile: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263-305). Mahwah, NJ: Lawrence Erlbaum.
Rogoff, B. (1995). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In J. V. Wertsch, P. d. Rio, & A. Alvarez (Eds.), *Sociocultural studies of mind* (pp. 139-164): Cambridge University Press.
Roschelle, J., Pea, R. D., Hoadley, C., Gordin, D. N., & Means, B. M. (2000). Changing how and what children learn in school with computer-based technologies. *The Future of Children, 10*(2), 76-101.
Sandoval, W. A. (2004). Developing learning theory by refining conjectures embodied in educational designs. *Educational Psychologist, 39*(4), 213-223.
Sandoval, W. A. (2013). Educational design research in the 21st century. In R. Luckin, S. Puntambekar, P. Goodyear, B. L. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of design in educational technology* (pp. 388-396). London: Routledge.
Sandoval, W. A. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences, 23*(1), 18-36. doi:10.1080/10508406.2013.778204

Sandoval, W. A. (2016). Disciplinary insights into the study of epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 184-194). New York: Routledge.

Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context [special issue]. *Educational Psychologist, 39*(4).

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction, 23*(1), 23-55.

Schwartz, D. L., Chang, J., & Martin, T. (2008). Instrumentation and innovation in design experiments: Taking the turn towards efficiency. In A. E. Kelly, R. A. Lesh, & J. Y. Baek (Eds.), *Handbook of design research methods in education*. New York: Routledge.

Shavelson, R. J., Phillips, D. C., Towne, L., & Feuer, M. J. (2003). On the science of educational design studies. *Educational Researcher, 32*(1), 25-28.

Steffe, L. P., & Thompson, P., W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 267-306). Mahwah, NJ: Lawrence Erlbaum Assoc.

Stromholt, S., & Bell, P. (2017). Designing for expansive science learning and identification across settings. *Cultural Studies of Science Education*.

Tabak, I. (2004). Reconstructing context: Negotiating the tension between exogenous and endogenous educational design. *Educational Psychologist, 39*(4), 225-233.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. London: Routledge.

This paper was commissioned for the Committee on Designing Citizen Science to Support Science Learning. The consensus study was convened by the Board on Science Education with support from Moore Foundation, Simons Foundation and Howard Hughes Medical Institute (HHMI). Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Science Education or the National Academy of Sciences, Engineering, and Medicine.