A Design-based Approach to Experimental Design: Investigating Hypotheses About How Framing Influences Transfer

Randi A. Engle, Sharla Roberts, Phi D. Nguyen, Pamela Yee and the Framing Transfer Research Group[1]
University of California at Berkeley, 4641 Tolman Hall 94720, RAEngle@berkeley.edu

Abstract: In this paper we present an approach to experimental design that is heavily influenced by practices from design-based research, illustrating it with our ongoing work to investigate hypotheses about how framing influences the transfer-of-learning. When one is interested in designing experiments that do not just test hypotheses about causal effects but also develop hypotheses about new kinds of causal mechanisms, we claim that it is particularly valuable to use the design-based research practices of iteratively designing new learning ecologies in particular contexts and creating explanatory accounts that coordinate multiple measures. Also, experimentally-inspired practices of comparing theoretically contrasting interventions in parallel and systematically streamlining designs to distinguish between critical and less critical design elements have promise for being usefully adapted within design-based research studies themselves.

Introduction

Randomized experiments are increasingly being advocated as a particularly strong way of conducting educational studies, with the US Institute for Education Sciences even prioritizing them for funding over other methodologies (Whitehurst, 2003). Advocates consider them to provide the most scientifically justifiable answers to questions from the public, policymakers, and practitioners about which educational interventions "work," which can support more evidence-based decisions in practice (e.g., NRC, 2002; Slavin, 2002; Whitehurst, 2003). In particular, random assignment of participants to interventions is said to support strong claims about the causal effects of particular interventions by creating comparison groups that are probabilistically equivalent to each other so that any differences in the outcome of one intervention compared to others can be attributed to the effect of the intervention (Shadish, Cook & Campbell, 2002).

However, not everyone in educational research is as sanguine about the merits of experimental methods. Although there are a wide range of objections (see Cook, 2007), here we focus on one. Namely, one key concern is that randomized experiments may provide reliable information that something "worked," but they often do not provide sufficient information about exactly what "it" was that worked nor "why" or "how" it worked (e.g., Erickson & Gutierrez, 2002; Maxwell, 2004; Olson, 2004). Even Shadish, Cook, and Campbell (2002), who generally argue for the power of experimental methods, state that:

> the unique strength of experimentation is describing the consequences attributable to deliberately varying a treatment. We call this *causal description.* In contrast, experiments do less well in clarifying the mechanisms through which and the conditions under which that causal relationship holds—what we call *causal explanation.* (Shadish, Cook & Campbell, 2002, p. 9)

Furthermore, it is in doing research that can support claims about the causal explanations for observed treatment effects that we can both construct stronger scientific theories as well as promote more successful applications in practice (e.g., Brown, 1992; Cook, 2007; Maxwell, 2004; Roschelle, Tatar & Kaput, in press).

At the same time, advocates of design-based research highlight its potential for strongly contributing to exactly this goal. For example, Cobb et al. (2003, p. 13) claim that in good design-based research,

> " 'What works' is underpinned by a concern for 'how, when, and why' it works, and by a detailed specification of what, exactly, 'it' is. This intimate relationship between the development of theory and the improvement of instructional design for bringing about new forms of learning is a hallmark of the design experiment [*a.k.a. "design-based research"*] methodology."

Similarly, the Design-based Research Collective (2003, p. 6) claim that design-based research studies can complement experimental ones "by assisting in the identification of relevant contextual factors, aiding in identification of mechanisms (not just relationships), and enriching our understanding of the nature of the intervention itself." More generally, both Maxwell's (2004) philosophically grounded defense of qualitative methods against narrow views of causality and Cook and colleagues' analyses of the merits and limitations of

experiments (e.g., Cook, 2007; Cook & Sinha, 2005; Shadish, Cook & Campbell, 2002) suggest that experimental studies can be significantly enhanced by incorporating other methods into them.

In this paper, we build on these and other suggestions to begin developing a design-based approach to experimental research. In particular, we show how several aspects of design-based research methods can crucially enhance the process of experimental design in ways that have the prospect of better contributing to both theory and practical application. We illustrate our approach by describing how we are designing a series of experiments in order to both test and further develop some new hypotheses about ways that transfer-of-learning can be affected by how learning contexts are framed interactionally (Engle, 2006).

First, we provide background on the topic and general design of the experiments we have been conducting. Next, we show how specific aspects of design-based methodology can contribute to experiments, especially those like ours that are focused on developing new causal explanations. Then, we discuss a few ways this approach to experimentation may also inform design-based studies. Finally, we close by briefly considering what may account for these resonances between what are otherwise two very different methodologies.

Background to the Experiments on Framing Transfer

The purpose of our research on framing transfer is to systematically develop and test a set of potentially influential causal mechanisms for fostering the transfer-of-learning that have been rarely investigated empirically. In particular, we are investigating the hypothesis that learning contexts that are interactionally framed as *expansive activities* in which learning and transfer contexts are intercontextually linked are more likely to support transfer-of-learning than those framed as *bounded events* (see Engle, 2006 for more). Bounded learning events are restricted to the here-and-now, include a narrowly defined set of topics and participants, and provide little space for learners' own ideas. In contrast, expansive learning activities extend across time and space, encompass a wide range of topics and participants, and incorporate learners' own ideas in an integral way. In general, we are suggesting that it is not just the content of what students learn that matters for whether they are likely to transfer what they have learned, but also how the contexts in which students learn that content are defined interactionally (e.g., Greeno, et al., 1993; Lab. of Comparative Human Cognition, 1983; Pea, 1987).

Consistent with Cobb and diSessa's (2004) account of how design-based research can support ontological theory building, we developed the hypotheses about framing learning contexts through empirical work within a design-based research project, Brown and Campione's Fostering Communities of Learners (FCL) (circa Brown & Campione, 1994). Inspired by Brown's (1992, p. 153) suggestion that researchers test "trends discovered in spontaneous classroom discussions…in the laboratory under more controlled conditions," we decided to design a series of experiments to see if we could identify a systematic effect of framing on transfer and, if so, better understand the underlying causal explanations for it. This would then provide a basis for later design-based research in classrooms by providing potential collaborating teachers with evidence that framing is effective as well as some initial ideas about how to implement it (Brown, 1992; McCandliss, Kalchman & Bryant, 2003).

We chose to conduct tutoring experiments to simultaneously address goals of simplicity, ecological validity, and theoretical relevance. Tutoring is both a common educational activity that our theory should be able to account for and one that allows for easier control and measurement given that only one student participates as a time. Each experiment is conducted over two sessions in which we first tutor students about the cardiovascular system by asking them to self-explain a text and diagrams (Chi, 2000; Chi et al., 2000). Self-explaining the same materials both makes is possible to control for content-based supports for transfer and provides a basis for the tutor to differentially frame the same student explanations as either reporting what the authors of the texts or diagrams say (bounded condition) or using these materials to author one's own ideas (expansive condition). We then assess the degree to which students transfer what they have learned to a related system, the respiratory system (Hmelo-Silver & Pfeiffer, 2004), using pre and post tests in both sessions to measure both learning and transfer. Matched pairs of students with similar initial knowledge about the systems are randomly assigned to two groups, one in which the sessions are framed as bounded events versus another in which they are framed as part of an expansive activity.

Contributions of Design-based Research to Experimental Design

In this section, we specify four aspects of design-based research (Cobb et al., 2003; Collins, Joseph & Bielaczyc, 2004; Design-based Research Collective, 2003; diSessa & Cobb, 2004) that can make particularly strong contributions to experiments like ours that seek to support claims about new and complex causal mechanisms:

1. Designing a learning ecology to specify and further develop a theory
2. Using iterative design and re-design cycles to better embody that theory in practice
3. Designing for particular educational contexts
4. Coordinating multiple measures to construct explanatory accounts

In each subsection below, we first discuss the merits of incorporating each of these design-based research aspects[(2)] within experiments and then illustrate them with our experiment about framing and transfer.

Designing a learning ecology to specify and further develop a theory

One of the hallmarks of design-based research is its ability to help researchers develop and flesh out theories (Cobb et al., 2003; Collins, et al., 2004; Design-based Research Collective, 2003; diSessa & Cobb, 2004). In embodying theories in practice, researchers become aware of vaguenesses, ambiguities, and other forms of under-specification in their theories that need to be addressed. The same occurs when designing experimental manipulations that seek to test a complex educational theory, especially one that endeavors to compare the efficacy not just of straightforward manipulations, but of different kinds of what Cobb and colleagues (2003, p. 9) refer to as "learning ecologies" (Cobb et al., 1993, p. 9), "complex, interacting system[s] involving multiple elements of different types and levels [that] function together to support learning."

The tutoring protocols our team designed for its experiments on transfer worked in similar ways to specify and further develop our emerging theory of how framing affects transfer. Consider the differences in the level of specification between Table 1, which is adapted from the paper that originally proposed our hypotheses (Engle, 2006, p. 491), and Table 2, which represents the current design elements that embody them. Table 1 presents fairly vague although theoretically-important generalities while Table 2 specifies particular actions that the tutors and study recruiters do in order to frame tutoring sessions as either expansive activities or bounded events. The ideas from Table 2 have been specified further in written protocols that the tutor uses to help create the larger learning ecology. This goes beyond simply operationalizing "variables" because what is embodied is a system of interacting elements that work together.

It is in the connections made between the general theoretical descriptions in Table 1 with their operationalizations in the particular design represented in Table 2 we have made our most theoretical progress. For example, we learned that at least three kinds of actions are consequential for framing time as bounded versus expansive: the tense and aspect of regularly used verbs, whether references are made to times besides the just completed present, and whether students are told they are participating in one study extending across both days (expansive) or two studies, one per day, that each consist of separate events (bounded). These actions build upon, but go beyond how the teacher in the original study framed time more expansively (see Engle, 2006, pp. 482-485). In addition, they specify what is involved in doing the opposite (that is, to frame time as bounded), a crucial step for making a clear comparison. Finally, this level of specification of the contrasts means we will know what to attribute effects to, something often hard to discern in randomized field trials.

Table 1: Hypotheses About Ways of Framing Contexts to Encourage or Discourage Transfer

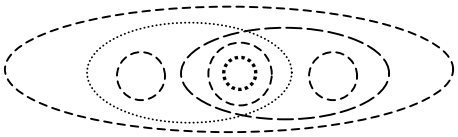
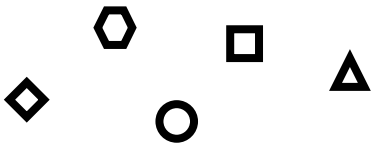
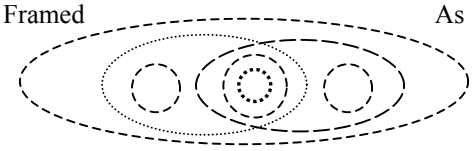
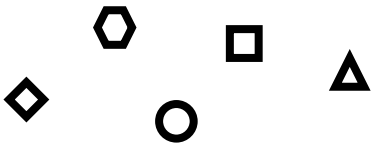
Aspects of Contexts That Can Be Framed	Framing As Expansive Activities  (Predicted to Promote Transfer)	Framing As Bounded Events  (Predicted to Discourage Transfer)
Setting • Time • Place • Participants	Broadly defined • Ongoing & connected • Broadly defined & extendable • Open, expanding set of participants, co-present and imagined	Narrowly defined • Set starting and ending points • Narrowly defined & circumscribed • Small, fixed set of participants
Topics	Connected, multiple topics	Disconnected, single topics
Roles	Authoring one's own ideas within activities framed broadly to include those ideas in an integral way	Reporting about others' ideas within an event framed narrowly as not including the learner's own ideas

Table 2. Operationalization of Framing Manipulations in the Tutoring Experiment

Aspects of Contexts That Can Be Framed	Framed As  Expansive Activities (Predicted to Promote Transfer)	Framed As Bounded Events  (Predicted to Discourage Transfer)
Setting: • Time • Place • Participants	Ask student to specify other settings in which the topic(s) have, are, or will be likely to come up in their lives • Refer to the study as a whole as including both days • Refer to other times, both inside and outside of the experiment • Use present progressive verbs ("you're figuring out") • Frame location as a university • Refer to other places—their home, school, doctor's office, etc.—in which they can use what they're learning • Treat larger activity as involving the student, you and the rest of the study team, plus their family, friends, teachers, and anyone else they mention above • Ask student how they would explain their ideas to other people besides you • When they show understanding of one of the key ideas, note that they can now explain that to whoever they mentioned as an audience, who will be pleased	Do not ask student to specify other settings in which the topic has, is, or will be likely to come up in their lives • Refer to each part of each day's session as a separate event • Make no references to times other than the just completed present • Use simple past with completion verbs ("we're finished with that now") • Frame location as this specific room • Make no references to other places • Treat tutoring event as a private matter involving only you and the student, and not other members of study team or other people they know • Only have student explain the text's ideas to you • When student shows understanding of one of the key ideas, simply praise them for properly reporting the text
Topics	• Both sessions together a single study • Whole study about body systems, with circulatory one of them	• Each session a separate study scheduled together for convenience • Today's study about the circulatory system
Roles	• Student asked to explain their own evolving ideas about the system using the text sentences as a resource • Revoice student's explanations, crediting student with authorship and checking with them about whether you reformulated their ideas accurately	• Student asked to explain what the text has said about the system in each sentence • Summarize what the text had presented

Unlike some versions of experimental methods but consistent with the design-based research notion of studying a learning ecology, we are in fact manipulating many related aspects of the framing and coordinating them with content-based aspects of the instruction (e.g., self-explaining) that are common across conditions. In so doing, we have a better chance of understanding how framing interacts with these instructional elements in the two different ecologies we create. This also provides a basis for later experiments to disentangle which aspects of the manipulation have which kinds of effects (see "Systematically streamlining…" below).

Using iterative design and re-design cycles to better embody a theory in practice

Actually moving from theoretical generalities to increasingly worked out contrasting interventions that can then be experimentally tested can be further supported by the design-based research strategy of using iterative design and re-design cycles (e.g., Cobb et al., 2003; Collins, et al., 2004; Design-based Research

Collective, 2003). Using this aspect of design-based research during an extended piloting stage while designing an experiment makes it possible to test and investigate more complex interventions.

Similarly, the progress we made in specifying the hypotheses about how framing affects transfer in the learning ecology of our experiment crucially relied on an iterative process of design and re-design during piloting. After each effort to embody the expansive activity or bounded event framing with a student, the tutor and the videographer, who had collected most of the data on the session, met to critique what just happened. This critique was based on a wide variety of data that had been just collected:

- a brief interview videographers conducted when the tutor was not present near the end of the session about the student's perceptions of their and the tutor's role in it,
- a short Likert-style questionnaire in which students were asked to rate their perceptions about how topics, settings, and roles were framed by the tutor and their degree of uptake of those framings,
- notes made by both videographer and tutor about any ways in which the tutor instantiated or departed from the tutoring protocol during the session, and
- notes made by the videographer about evidence related to the student's uptake (or lack thereof) of the framing manipulation during the tutoring session (later converted into an observational checklist)

The goal of this critique was to identify any aspects of what the tutor did that may have unwittingly:

- provided a substantive advantage to one condition or the other in terms of content-based supports for transfer rather than keeping them constant across conditions
- in some way muddied the intended contrast between the conditions by embodying aspects of an expansive activity framing within the bounded event condition or vice versa

We then adjusted the protocols accordingly to more explicitly direct the tutor's actions in ways that would strengthen the manipulation while holding other aspects of the intervention constant.

So although there is iterative design and re-design process in creating an experiment like this one, it is targeted towards a different goal than in a typical design-based research study. Rather than trying to tune an intervention so it increasingly results in the best outcomes for students through embodying a particular theory, the focus of designing and re-designing in the case of a theory-based experiment like ours is to systematically embody just the theory-relevant hypotheses that we wish to test *without* optimizing other aspects of the design that could have given undue advantage to one condition over a contrasting one. However, by collecting and reflecting on a wide variety of data sources about the nature of the implementation as in a design-based research study, it is possible to refine even a complex experimental instructional protocol so it is increasingly more representative of the theoretical contrasts one wishes to test. Although one could say that this is just good piloting, the few texts on experimental methods that mention piloting tend to ignore theory-relevant issues, focusing on practicing procedures and designing reliable and sensitive measures (e.g., Davis & Rose, 2000; Leech, 1991).

One open question, however, is the status of the resulting data as an experiment. If over the course of an experiment, the design gets increasingly refined so it is more and more tightly controlled, with the contrast more closely reflective of the theoretical contrasts of interest, what does it mean if one gets a statistically reliable effect in outcomes? Although this clearly violates experimental maxims to keep treatments fixed over the course of an experiment, we suggest that finding an effect under such conditions would be supportive of a hypothesis (especially if it got stronger from student to student as we are currently observing). Concerns with variability of treatments then could be addressed by replicating an experiment with a fixed protocol after sufficient iterations.

Designing for particular contexts

Another feature of design-based research that can be incredibly valuable for experimental design is the idea of viewing an experimental session in the laboratory as creating its own unique learning context with participants who come from their own particular learning contexts (Cobb et al., 2003; Collins et al., 2004; Design-based Research Collective, 2003). Again, by context, we are referring to each situation as interactionally define by participants: what it is they are doing together and who they are being in doing that. As we will illustrate below, by considering context within the purview of experimental design we can both better understand the contributions of context on experimental outcomes while better judging where our findings are likely to be most applicable.

In the framing transfer experiment, context has been important from the start as it is the context of the experimental sessions itself that we are manipulating in order to understand how it affects the nature of students' transfer. At the same time, we work to make the experimental sessions similar to other educational contexts by using instructional materials comparable to those that students use in schools. We also record any references students make to other contexts during learning to identify examples of students creating intercontextual links between contexts within and outside of the experiment.

In addition, we have increasingly been designing our experiment for students from particular pre-existing educational contexts. At first, we recruited students from a wide range of contexts and discovered that they seemed to interpret our framing manipulations differently depending on how their prior learning contexts

had been framed. Because of this, we are now recruiting students from the same biology classes within the same schools. This allows us to observe how their biology instruction is being framed and adjust our framing manipulation accordingly so it will be more likely to be taken up by them in the ways we intend. In so doing, we reduce within-sample variability in students' propensities to take up different framings. We also can better tune the substantive aspects of our instruction across conditions to build on students' likely prior knowledge.

Finally, designing in particular contexts can suggest additional non-experimental contrasts for investigation. In our case, we have decided in future studies to systematically vary the student populations that participate to include some who had experienced at least some aspects of the expansive activity framing before. These students would contrast with our current crop of students whose previous classroom experiences are in fact even more bounded than our current bounded event condition as their primary learning activities are reading textbooks and listening to lectures without any audience (like the tutor) to listen to them as they make sense of what these materials say. Our prediction is that student populations who have experienced both framings will be able to take them up more readily than those who have only experienced one, and will those show greater difference in transfer.

Coordinating multiple measures to construct explanatory accounts

Design-based research studies collect multiple measures of a design's ecology and its potential outcomes to: identify the kinds of outcomes a design actually has; triangulate measures so findings are more reliable; provide data that could conflict with theoretical expectations; and create sophisticated explanations of how a design works that relate aspects of its ecology to outcomes (Cobb et al., 2003; Collins et al., 2004; Design-based Research Collective, 2003). Similarly, experiments designed to develop causal explanations, especially if they involve learning ecologies that have not been studied extensively, can profit considerably from similar densities of data collection and analysis.

In the case of the framing transfer experiment, the multiple measures of implementation mentioned above for refining the design have been supplemented by multiple audio and video records of each tutoring or transfer of learning session. These records provide opportunities to revisit hypotheses (Engle, Conant & Greeno, 2007) about how students responded to each framing, and how those responses related to potential mediating mechanisms, including students' choices of where to focus their attention and when to bring in prior knowledge. This will support much more well-grounded explanations for the effects we observe.

In addition, as the prior literature does not agree on a single operational definition for transfer and there are a wide variety of types of knowledge transfer that framing might affect, we purposely designed the experiment so multiple types of transfer were both possible and measurable. In particular, we assess transfer of knowledge that ranges in complexity from basic terminology to conceptual principles to learning practices. At the same time, we measure transfer at multiple time points during the experiment: immediate transfer on day 1 to near transfer tasks, initial performance-based transfer to the respiratory system during its pre-test on day 2, references to information about the cardiovascular system while learning about the respiratory system on day 2, and "preparation-for future learning" style facilitation effects as measured by differences in pre/post gains in learning about the respiratory system (Bransford & Schwartz, 1999). Using multiple measures of transfer will allow us to identify an effect of framing on transfer if one is to be found. At the same time, it will allow us to learn whether framing affects different kinds of transfer in different ways. If so, we will then need to develop our theory further to account for such differential effects. Had we measured just one outcome variable for transfer, none of this would be possible. In general, we agree with Cobb et al. (2003, p. 12) that "multiple sources of data ensure that retrospective analyses conducted when the experiment has been completed will result in rigorous, empirically grounded claims and assertions."

Potential Contributions to Design-based Research Studies

In parallel with how design-based research principles have been useful for designing rich experiments that can better support the testing and development of emerging theories, there are a few aspects of our hybrid approach that could be usefully adapted for design-based research studies themselves.

The parallel design of systematically contrasting interventions

First is the potential power of engaging in a design-based research project in which one designs systematically contrasting interventions in parallel, which is similar to what we have been doing by framing the same instructional design in two contrasting ways. For ethical reasons, one would not want to do this in schools when one had good a priori reasons to believe that one manipulation would likely be superior to another. However, this could be particularly fruitful when a team is faced with a design choice between two options, each of which appears to have both advantages and disadvantages. Comparing them with each other using crossed or split lagged designs across multiple classes taught by the same teacher could be a effective methodology for both learning about the actual impacts of the design choice and advancing theoretical claims

relevant to it. In so doing we expand on Brown (1992) and others' use of full and partial control groups within design-based research projects by promoting theory-based comparisons as well.

Systematically streamlining interventions to identify necessary and unnecessary elements

At one point in Ann Brown's classic paper on design experiments, she notes that she needs to find some way to "unconfound variables, not only for theoretical clarity, but also so that necessary and sufficient aspects of the intervention can be disseminated" (Brown, 1992, p. 173). In our work, as mentioned above, we purposely manipulated a whole host of aspects of framing that we hypothesize will collectively produce an expansive activity versus bounded event framing. However, if we find an effect of this framing, then our next step will be to pursue a series of follow-up experiments in which we systematically reduce the manipulation so that fewer frameable aspects of the learning context (i.e., roles, topics, and settings, and within settings, time, place, and participants) are being manipulated at once. In so doing, we will be able to better learn which aspects of framing matter the most and how they might interact with each other. Similarly one can begin identifying the necessary aspects of a design ecology by systematically streamlining it by eliminating features hypothesized to be unnecessary, and then comparing to findings found with the original version (cf. Kelly, 2004).

Discussion: Explaining the Resonances Between the Methodologies

What explains the resonances we have identified between what are otherwise two radically different methodologies—our use of randomized experiments and design-based research (Collins et al., 2004)? Although undoubtedly some are due to general principles for empirically supporting arguments, two key shared research goals are also probably important: investigating new territory and developing causal explanations. A unique strength of design-based research is that it allows researchers to study innovative learning environments that otherwise would not have existed (Brown, 1992; Cobb et al., 2003; Design-based Research Collective, 2003). Similarly, in harnessing the power of design, experiments can be conducted that address a wider range of phenomena and hypotheses, including those still in development. In addition, when the goal is to investigate causal explanations, it makes sense for experiments as well as design-based research studies to provide "extensive measurement and analysis of theory-derived mediating processes" (Cook & Sinha, 2005, p. 558). At the same time, it makes sense for design-based research studies to also use systematic contrasts whenever possible to decide between competing explanations (e.g., Maxwell, 2004).

Endnotes

(1) The Framing Transfer Research Group also includes Pauline Huang who helped initiate the study and Adam Mendelson and Amy Stornaiuolo who provided suggestions and criticism on the paper. We also appreciate discussions around these issues with Dor Abrahamson, Cynthia Coburn, David Hammer, Roy Pea, Jeremy Roschelle, and Dan Schwartz as well as the reviewers' incisive feedback. Although all made contributions, only we are responsible for the weaknesses.

(2) To forestall misunderstanding, we will note that there are other aspects of design-based research that do not apply as readily to this kind of experimental research. For example, there is no particular requirement that one extensively collaborate with practitioners (e.g., Collins et al., 2004; Design-based Research Collective, 2003), create designs that are or would be considered to be particularly innovative (e.g., Collins et al., 2004; Kelly, 2004), or end up with a design product useful in a wide range of other settings (e.g., Kelly, 2004; Roschelle, Tatar & Kaput, in press).

References

Bransford, J.D., and Schwartz, D.L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*, 61-100.

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

Brown, A. L. & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-270). Cambridge, MA: MIT Press.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161-238). Mahwah, NJ: Erlbaum.

Chi, M. T. H., Siler, S., Jeong, H., Yamauchi, T., Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science, 25,* 471-533.

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

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

Cook, T. D. (2007). Randomized experiments in education: Assessing the objects to doing them. *Economics of*

Innovation and New Technology, 16(5), 331-355.
Cook, T. & Sinha, V. (2005). Randomized experiments in educational research. In J. L Green, Camilli, G. and P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 551-566). Mahwah, NJ: Erlbaum.
Davis, A. & Rose, D. (2000). The experimental method in psychology. In G. M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.), *Research methods in psychology* (pp. 42-58). 2nd Edition. London: Sage Publications.
Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–9.
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.
Engle, R. A. (2006). Framing interactions to foster generative learning: A situative account of transfer in a community of learners classroom. *Journal of the Learning Sciences, 15(4),* 451-498.
Erickson, F. & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher, 31*(8), 21-24.
Greeno, J. G., Smith, D. R., & Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial* (pp. 99–127). Norwood, NJ: Ablex.
Hmelo-Silver, C. E. & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science, 28,* 127-138.
Kelly, A. E. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences, 13*(1), 115-128.
Laboratory of Comparative Human Cognition. (1983). Culture and cognitive development. In P. H. Mussen (Ed.), *Handbook of child psychology: Volume 1. History, theory and methods* (pp. 295–356). New York: Wiley.
Leach, J. (1991). *Running applied psychology experiments.* Milton Keynes, UK: Open University Press.
Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11.
McCandliss, B. D., Kalchman, M. & Bryant, P. (2003). Design experiments and laboratory approaches to learning: Steps towards collaborative exchange. *Educational Researcher, 32*(1), 14-16.
National Research Council (2002). *Scientific research in education*. Washington DC: National Academies Press.
Olson, D. R. (2004). The triumph of hope over experience in the search for "what works": A response to Slavin. *Educational Researcher, 33*(1), 24-26.
Pea, R. D. (1987). Socializing the knowledge transfer problem. *International Journal of Educational Research, 11*, 639–663.
Roschelle, J., Tatar, D. & Kaput, J (in press). Getting to scale with innovations that deeply restructure how students come to know mathematics. In A. E. Kelly & R. Lesh (Eds.), *Handbook of innovative design research in science, technology, engineering, mathematics education.* Hillsdale, NJ: Erlbaum.
Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton-Mifflin.
Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher, 31*(7), 15-21.
Whitehurst, G. R. (2003, April). *The Institute of Education Sciences: New wine, new bottles*. Talk presented at AERA. Available at: http://ies.ed.gov/director/speeches2003/04_22/2003_04_22.asp.