THE JOURNAL OF THE LEARNING SCIENCES, *13*(1), 15–42

Design Research: Theoretical and Methodological Issues

Allan Collins
School of Education and Social Policy
Northwestern University

Diana Joseph
Center for School Improvement
University of Chicago

Katerine Bielaczyc
Harvard Graduate School of Education
Harvard University

The term “design experiments” was introduced in 1992, in articles by Ann Brown (1992) and Allan Collins (1992). Design experiments were developed as a way to carry out formative research to test and refine educational designs based on principles derived from prior research. More recently the term design research has been applied to this kind of work. In this article, we outline the goals of design research and how it is related to other methodologies. We illustrate how design research is carried out with two very different examples. And we provide guidelines for how design research can best be carried out in the future.

In the 1990s there has been a movement to develop a new methodology for carrying out studies of educational interventions under the labels “design experiments” or “design research,” which will be used interchangeably in this article. Ann Brown (1992) was a leader in this movement and this article is an attempt to carry forward her work to specify for the educational research community the basis for this movement and the research methods it entails.

Correspondence and requests for reprints should be sent to Allan Collins, 135 Cedar Street, Lexington, MA 02421. E-mail: collins@bbn.com

Design research was developed to address several issues central to the study of learning, including the following:

- The need to address theoretical questions about the nature of learning in context.
- The need for approaches to the study of learning phenomena in the real world rather than the laboratory.
- The need to go beyond narrow measures of learning.
- The need to derive research findings from formative evaluation.

Although design research is a powerful tool for addressing these needs, this kind of work brings with it serious challenges, including the following:

- Difficulties arising from the complexity of real-world situations and their resistance to experimental control.
- Large amounts of data arising from a need to combine ethnographic and quantitative analysis.
- Comparing across designs.

The notion of design research is very much in vogue in the learning-sciences community, but faces obstacles in the broader research community that make publication and tenure difficult for learning scientists. Whereas some of these obstacles arise simply from introducing a new methodological approach to the research world, some might be mitigated by actions we take as a field. If design research is to become accepted as a serious scholarly endeavor, the learning-sciences community needs to take responsibility for creating standards that make design experiments recognizable and accessible to other researchers, and for developing an infrastructure that can support summative evaluation.

The first section of this article describes where design research falls within the landscape of experimental methods. This section characterizes design research as a particular kind of formative research, and discusses what it can tell you and what it cannot tell you. The second section describes two examples of design research to show the range of methods and scales involved, in particular: Brown and Campione's work in the Oakland Schools around the Fostering a Community of Learners project, and the work of Diana Joseph in designing a passion curriculum in an afterschool setting at a Chicago school. The third section addresses questions about the conventional elements of a design experiment, including the variables to consider, the methods to use, and how to report findings. Finally, the fourth section discusses how the methods of design research might be applied to summative research.

INTRODUCTION

Design Sciences Versus Analytic Sciences

Herbert Simon (1969) in his classic book *The Sciences of the Artificial* makes a distinction between natural sciences and sciences of the artificial, by which he means design sciences. He identifies the various professions, such as architecture, engineering, computer science, medicine, and education, with the sciences of the artificial, and outlines what a curriculum in the engineering sciences might include (e.g., utility theory, control theory, etc.). In Simon's view the design sciences have been neglected, because of the lack of rigorous theories. He argues that recent developments in engineering and computer science have begun to provide the theoretical underpinnings that the sciences of the artificial need. Although he identified critical bodies of theory for the engineering sciences, his analysis does not provide the theoretical foundations for a design science of education.

In one of the first articles on design experiments, Collins (1992) framed his work as heading toward a design science of education. He discussed similar issues to those raised by Simon in terms of a distinction between analytic (or natural) sciences and design sciences. He viewed physics, biology, and anthropology as analytic sciences, where the effort is to understand how phenomena in the world can be explained. He viewed aeronautics, artificial intelligence, and acoustics as design sciences, where the goal is to determine how designed artifacts (e.g., airplanes, robots, or concert halls) behave under different conditions. Just as in aeronautics, where researchers look at how different designs affect lift, drag, and other dependent variables, he argued that we need to develop a design science of education, where we investigate how different learning-environment designs affect dependent variables in teaching and learning.

Design As Intention Versus Design As Implementation

There is a fundamental challenge in developing a design science of education in that the enacted design is often quite different from what the designers intended. Brown and Campione (1996) referred to this problem in terms of "lethal mutations," where the goals and principles underlying the design are undermined by the way the design is enacted. More generally, any implementation of a design requires many decisions that go beyond the design itself. This occurs because no design can specify all the details, and because the actions of participants in the implementation (e.g., students, parents, teachers, and administrators) require constant decisions about how to proceed at every level. Designs in education can be more or less specific, but can never be completely specified. Evaluation of designs can only be made in terms of particular implementations, and these can vary widely depending on the participants' needs, interests, abilities, interpretations, interactions, and goals.

Hence in evaluating any design, it is important to keep in mind the limitations of the evaluation. The effectiveness of a design in one setting is no guarantee of its effectiveness in other settings. This is a fundamental problem pervading all education research, which we return to later when we discuss the limitations of different methodologies for carrying out education research.

The Problem of Narrow Measures

A related problem that research in education faces is the fact that most experiments treat a single variable, such as learning of content and skills, as the only measure of worth. This problem reveals itself most clearly in the notion that a multiple-choice test of content and skills can be used as the "bottom line" measure of how a student, teacher, or school is doing. In fact, employers care much more about a students' dispositions than about the particular content and skills that students acquire in school (Rosenbaum, 2001; Stasz, 2001). They care whether potential employees show up on time, take responsibility for getting a job done, work well with others, put in their best effort to succeed, and so forth. They would probably regard a school that instilled these virtues as much more valuable than a school that simply instilled a lot of specific knowledge. A similar argument can be made that these dispositions are key to success in college and any other endeavor that students are likely to undertake.

But this is just one aspect of the narrowness with which we evaluate interventions. A central concern of teachers is whether they motivate their students to want to learn. If students come out of school as proficient test takers who hate school, math, or Shakespeare, then the school will have failed. Education needs to produce "expert learners" (Brown, Ellery, & Campione, 1998), who love learning and who know how to find things out for themselves. Educators need to create environments where students are not afraid to put forth new ideas, share what they learn, and produce products they can show to the world. If we do not evaluate our educational environments in these terms then we will misjudge them badly and worse, we will lead them to emphasize the wrong goals (Frederiksen & Collins, 1989; Shepard, 2000).

Characteristics of Design Research

Design experiments were developed as a way to carry out formative research to test and refine educational designs based on theoretical principles derived from prior research. This approach of *progressive refinement* in design involves putting a first version of a design into the world to see how it works. Then, the design is constantly revised based on experience, until all the bugs are worked out. Progressive refinement in the car industry was pioneered by the Japanese, who unlike American car manufacturers, would update their designs frequently, rather than waiting years for a model changeover to improve upon past designs. The approach also is the basis of Japanese lesson study (Stigler & Hiebert, 1999) where groups of teachers meet together to refine their teaching practices. By studying a design in practice with an eye toward progressive refinement, it is possible to develop more robust designs over time. In a later section of the article we discuss how the design-research methodology might be applied to summative research.

Design research is not aimed simply at refining practice. It should also address theoretical questions and issues if it is to be effective. For example, Diana Joseph (2000), in work described later in this article, assessed a motivational theory developed with Daniel Edelson (Edelson & Joseph, in revision; Joseph & Edelson, 2002). She analyzed the motivational patterns among different participants and how they played out in their work in an afterschool setting. This analysis led to refinements in the design, but also fostered refinements in the theory. Design research should always have the dual goals of refining both theory and practice.

Because design experiments are set in learning environments, there are many variables that cannot be controlled. Instead, design researchers try to optimize as much of the design as possible and to observe carefully how the different elements are working out. Such observation entails both qualitative and quantitative observations, just as Consumer Reports evaluates products in terms of both qualitative and quantitative measures. When some aspect of the design is not working, the design team, including the teacher or facilitator, should consider different options to improve the design in practice, and institute design changes as frequently as necessary. Any changes to one aspect of the design need to be considered with respect to how well they fit with other aspects of the design, since as Brown and Campione (1996) have emphasized, the design needs to be thought of as an integrated system. Thus the evaluation of the design is an ongoing process that changes as the design changes.

Design experiments have some fundamental limitations. Because they are carried out in the messy situations of actual learning environments, such as classrooms or afterschool settings, there are many variables that affect the success of the design, and many of these variables cannot be controlled. Design researchers usually end up collecting large amounts of data, such as video records of the intervention and outputs of the students' work, in order to understand what is happening in detail. Hence, they usually are swamped with data, and given the data reduction problems, there is usually not enough time or resources to analyze much of the data collected. It also takes resources to collect so much data, and so design experiments tend to be large endeavors with many different participants, all of whose work needs to be coordinated. All these factors make design experiments difficult to carry out and the conclusions uncertain.

Characteristics of Other Research Methods

All methodologies have their advantages and disadvantages. Despite the limitations of design experiments, they have an important role to play among the different methodologies for studying educational practice. In order to place design experiments in the landscape of different methods we will compare them to three general types of educational evaluation approaches.

Laboratory and training studies attempt to control variables in order to determine how particular independent variables affect a few dependent variables such as the learning of content and skills. Ann Brown (1992) felt that training studies neglect important variables that affect the success of a design. In her view they also tended to be carried on over such a short period of time that there was no way they could identify what were the major effects of the design on learning.

Collins (1999) compared laboratory studies of learning to design experiments in terms of seven contrasting aspects of their methodology:

1. *Laboratory settings vs. messy situations.* Experiments conducted in laboratories avoid contaminating effects. Learners concentrate on the task without any distractions or interruptions. The materials to be learned are well defined and are presented in a standardized manner. Design experiments are set in the messy situations that characterize real life learning, in order to avoid the distortions of laboratory experiments.

2. *A single dependent variable vs. multiple dependent variables*. In most psychological experiments there is one dependent variable, such as the number of items recalled or the percent correct on a test of some kind. In design experiments there are many dependent variables that matter, though the researchers may not pay attention to them all.

3. *Controlling variables vs. characterizing the situation.* Psychological experiments use a methodology of controlling variables, where the goal is to identify a few independent and dependent variables, and hold all the other variables in the situation constant. In design experiments, there is no attempt to hold variables constant, but instead the goal is to identify all the variables, or characteristics of the situation, that affect any dependent variables of interest.

4. *Fixed procedures vs. flexible design revision*. Psychological experiments follow a fixed procedure that is carefully documented, so that it can be replicated by other experimenters. Design experiments, in contrast, start with planned procedures and materials, which are not completely defined, and which are revised depending on their success in practice.

5. *Social isolation vs. social interaction*. In most psychological experiments, the subjects are learning in isolation. There is no interaction with other learners and usually no interaction with a teacher or expert; the material to be learned is simply presented by text or video. By contrast, design experiments are set in complex social situations, such as a classroom.

6. *Testing hypotheses vs. developing a profile*. In psychological experiments the experimenter has one or more hypotheses, which are being tested by systematically varying the conditions of learning. In design experiments the goal is to look at

many different aspects of the design and develop a qualitative and quantitative profile that characterizes the design in practice.

7. *Experimenter vs. co-participant design and analysis*. In psychological experiments the experimenter makes all decisions about the design and analysis of the data, in order to maintain control of what happens and how it is analyzed. In design experiments, there is an effort to involve different participants in the design, in order to bring their different expertise into producing and analyzing the design.

Laboratory studies are effective for identifying effects of particular variables, but they often neglect variables critical to the success of any intervention. Ann Brown (1992) valued them for their role in developing a design, but to test and refine the design requires going into complex settings.

Ethnographic research also is set in the contexts of natural learning environments, but it has some of the characteristics of the analytic sciences, in that it does not attempt to study interventions. In general, ethnographic research (e.g., Eckert, 1989) attempts to characterize relationships and events that occur in different educational settings. There is no attempt to change educational practice, as in design experiments, but ethnographic research produces rich descriptions that make it possible to understand what is happening and why.

Large-scale studies of educational interventions use a variety of measures to determine the effects of a program or intervention. The methods usually emphasize standardized measures and surveys of critical participants, not tied to any particular design. These studies can be used to identify critical variables and to evaluate program effectiveness in terms of test scores, but they do not provide the kind of detailed picture needed to guide the refinement of a design. They are crucial however for summative research, and in our discussion of how design experiments methodology might be extended to summative research, we borrow from this kind of methodology.

Why Design Research?

Design experiments bring together two critical pieces in order to guide us to better educational refinement: a design focus and assessment of critical design elements. Ethnography provides qualitative methods for looking carefully at how a design plays out in practice, and how social and contextual variables interact with cognitive variables. Large-scale studies provide quantitative methods for evaluating the effects of independent variables on the dependent variables. Design experiments are contextualized in educational settings, but with a focus on generalizing from those settings to guide the design process. They fill a niche in the array of experimental methods that is needed to improve educational practices.

EXAMPLES OF DESIGN RESEARCH

To show some of the variety of design experiments carried out in recent years, we will describe two case studies. The first case is Ann Brown and Joseph Campione's Fostering a Community of Learners (FCL), which was carried out in a variety of elementary schools across the country. This illustrates design research that studies how a classroom intervention plays out in a number of classrooms at different levels in an elementary school. The second case we describe is Diana Joseph's development and investigation of the passion curriculum approach, a framework for developing interest-centered learning environments. This work was conducted through a video-making club, which served as a prototype passion curriculum, enacted in an urban public school classroom and afterschool program. This case illustrates a small-scale design experiment, where a single researcher conducted a study in order to refine the design parameters for a new type of curriculum.

Brown and Campione's FCL Classrooms

Brown and Campione (1994, 1996; Brown, 1992) developed a model they call Fostering a Community of Learners (FCL) for grades 1–8. The model provides what is termed a "developmental corridor," where the learning community extends not only horizontally across a classroom, but also vertically across grades. This makes it possible for learning topics to be revisited at increasing levels of disciplinary sophistication.

The FCL approach promotes a diversity of interests and talents, in order to enrich the knowledge base of the classroom community as a whole. The focus of FCL classrooms is on the subject areas of biology and ecology, with central topics such as endangered species and food chains and webs. There is an overall structure of students (1) carrying out research on the central topics in small groups where each student specializes in a particular subtopic area, (2) sharing what they learn with other students in their research group and in other groups, and (3) preparing for and participating in some "consequential task" (Scardamalia, Bereiter, & Fillion, 1981) that requires students to combine their individual learning, so that all members in the group come to a deeper understanding of the main topic and subtopics. Teachers orchestrate students' work, and support students when they need help.

There are three research cycles per year. A cycle begins with a set of shared activities and materials meant to build a common knowledge base. Students then break into research groups that focus on a specific research topic related to the central topic. For example, if the class is studying food chains, then the class may break into five or six research groups that each focus on a specific aspect of food chains, such as photosynthesis, consumers, energy exchange, and so forth. Students research their subtopic as a group, with individuals "majoring" by following their own research agendas within the limits of the subtopic. Students also engage in “crosstalk,” talking across subtopic groups to explain, ask questions, and refine their understanding. The research activities include reciprocal teaching (Palincsar & Brown, 1984), guided writing and composing, consultation with subject matter experts outside the classroom, and cross-age tutoring. In the final part of the cycle, a member from each of the subtopic groups come together to form a “jigsaw” group (Aronson, 1978) in order to share learning on the various subtopics and to work together on some consequential task. Thus, in the jigsaw, all pieces of the puzzle come together to form a complete understanding.

The consequential task requires the different subtopics to be used together to form a common product or common understanding. The choice of consequential tasks is ideally made by the teacher and students together. The consequential task might be a bulletin board display, the design of a bio-park to protect an endangered species, a presentation to the community at large, or in some cases a test of students’ knowledge. These tasks “bring the research cycle to an end, force students to share knowledge across groups, and act as occasions for exhibition and reflection” (Brown & Campione, 1996, p. 303).

We can trace the evolution of FCL as a series of design experiments that culminated in the design described above. The story of the evolution of FCL is a story about successive refinements of a design, so that it better addressed the goals and principles that Brown and Campione (1994, 1996) have developed. As Campione (2000) made clear in the Scribner Award talk at the American Educational Research Association meeting, he and Ann Brown saw their work as going from laboratory research to learning principles to the design of a learning environment. Then based on their analysis of the learning environment, they would make modifications and additions to the learning principles, which in turn led to modifications of the learning environment and new laboratory experiments.

We begin the story of the evolution of FCL with the experiments by Palincsar and Brown (1984) on “reciprocal teaching” of reading. In reciprocal teaching of reading, students are expected to

1. Ask a question about the text.
2. Identify anything that is unclear.
3. Summarize what they have just read.
4. Make a prediction about what is coming next, after each paragraph of text.

The first experiments were done with individual children in the laboratory, and showed enormous gains in their ability to answer questions on reading comprehension tests. Given a procedure that was so successful with individual children, Palincsar and Brown extended the design to groups of students in classroom settings, since this is how reading is generally taught in American schools. To do this they had each child take a turn at being “teacher,” where they asked other students to carry out the four reading comprehension tasks listed previously. Again there were very large gains in reading comprehension scores for the groups of students engaged in reciprocal teaching.

Phase 1. Brown and Campione (personal communication) found that reciprocal teaching was limited in that the student discussions never were very deep. The discussions were restricted to the text at hand, and so the students tended not to bring in knowledge from other sources. Brown and Campione wanted to create a context where students would be integrating knowledge across multiple texts. In such a context students would build theories about a domain that they could bring to bear in their discussions. It was this attempt to enrich the content of what students were discussing that led to the design of FCL.

In the first implementation of FCL Brown and Campione had students study interdependence and adaptation in biological systems, working in three Urbana Illinois classrooms. Reciprocal teaching was woven throughout the design, whenever students were reading materials. The students were broken into groups to study different aspects of interdependence and adaptation, such as camouflage, animal communication, and predator-prey relationships. The students studied books and other materials the project provided, and then they wrote up what they learned on computers. When each research group had prepared a document describing what they had learned, the students broke up into jigsaw groups, where one student from each of the research groups met with students from the other groups. That student would play the role of the teacher when the jigsaw group read the material that his or her research group had written. The task of each jigsaw group was to put together a booklet on biological systems, which was comprehensible to everyone in the group. In this first version of FCL there was a new emphasis on students gaining deeper understanding by writing explanations for other students and engaging in jigsaw with other students, where each shared their knowledge with students who had worked on other topics.

There were two important theoretical ideas that emerged out of this first phase of FCL: the notion of *diverse expertise* and the notion of a *community of learners*. The initial goal of the design was to encourage students to deeply explore content and to develop their own understanding by generating explanations of what they had learned, both orally and in writing. But Brown and Campione found that more important social goals had also been achieved by the design. Students came to value the expertise of other students; not just content expertise, but sometimes expertise in using computers or in keeping the group working effectively toward their goal. It became clear that students worked together better when they appreciated others' contributions. This is how the idea of diverse expertise took hold, with its emphasis on respect and listening to others. From this developed the view of a community of learners, which is radically different from the emphasis on individual learning that permeates schooling. The idea of a community of learners emphasizes students working to understand different aspects of a domain and sharing their knowledge with everyone, rather than the traditional notion of all students learning the same thing at the same time.

Phase 2. Despite the initial success of the FCL design, Brown and Campione found from the work students produced that there were many misconceptions about biological systems, which the students were developing. For example, their ideas about evolution were mostly Lamarckian rather than Darwinian. Hence, when Brown and Campione moved to California, they revised the design of FCL to put more emphasis on biological content. To do this they added benchmark lessons and hands-on activities. In the benchmark lessons at the beginning of a unit, a teacher or outside expert introduced key ideas about biology in order to provide the students with the background for their research. At the same time they introduced laboratory activities where students could carry out hands-on experiments related to the topics they were researching. This dramatically expanded the kind of research that the students could carry out.

In this setting computers were still only being used as word processors, which was a waste of a major resource. Hence, Brown and Campione further redesigned FCL to have students go out on the Web in order to find information relevant to their research topics. They also set up telementoring relationships with biology experts (Brown, Ellery, & Campione, 1998). This gave the students access to real expertise from outside the classroom and acted to validate the importance of their research to the students. For example, one class of students became interested in the question of whether AIDS could be transmitted by mosquitoes, and so they investigated this topic thoroughly, calling upon resources and experts from outside the classroom, since the adults in the classroom did not have the information they were seeking. Students clearly gained the ability to address challenging questions of their own making.

These changes extended the notion of community beyond the classroom in order to bring in more expertise from outside the classroom: they were developing a wider *community of practice*, where students had *multiple ways in* to participate. They were no longer limited by the ideas within the classroom, but could go outside to address questions that arose from their research. At the same time Brown and Campione found that the hands-on activities did not integrate well with the other research activities the students were carrying out, and so they decided to discontinue the hands-on activities. Out of this phase they also developed the notion of a *developmental corridor,* where students would cycle through related topics in biology over the years of elementary school, but each time in greater depth, reflecting the developmental level that students were ready to achieve. This last idea required moving to a new elementary school in Oakland, where teachers at different grade levels were willing to participate in implementing FCL in their classrooms.

Phase 3. In the third phase of the development of FCL, Brown and Campione worked with teachers from second to eighth grade among a largely minority population. This allowed then to implement a developmental corridor that led children "gradually toward deep principles of the discipline, such as interdependence, biodiversity, evolution, and adaptation" (Brown & Campione, 1996, p. 306). "By second grade we begin to address animal/habitat mutuality and interdependence. Sixth graders examine biodiversity and the effect of broad versus narrow niches on endangerment. By eighth grade the effect of variation in the gene pool on adaptation and survival is not too complex a topic." (Brown & Campione, 1996, pp. 307–308). Implementing the developmental corridor also allowed them to set up cross-age tutoring where students in the higher grades tutored second and third graders on the topics they were investigating (e.g., see Brown, Ellery, & Campione, 1998).

A final refinement to the FCL design came from the students themselves. They pointed out that the formal sharing of knowledge came at the end, when the jigsaw groups got together. The students suggested therefore that they needed to have "*crosstalk*" sessions, where students from the different research groups present their intermediate findings to the whole class. Sometimes probing questions from other students leads students down new research paths to resolve the questions that arise during the crosstalk. It is also a forum for raising issues that the groups are worried about. Crosstalk has become a major vehicle for sharing knowledge and building a learning community.

This third phase of their work put together all of the pieces that make up the FCL design outlined in Brown and Campione's 1996 article. Table 1, taken from Brown and Campione (1996), captures the design principles that evolved from the various phases of the FCL design, together with their previous work on metacognition, analogy, dynamic assessment, the zone of proximal development, and reciprocal teaching. The FCL design, in fact, brings together 30 years of development work as to how best to structure a learning environment for students in schools.

Brown and Campione (1994, 1996; Brown, 1992) have collected a wide variety of data about students and teachers in the FCL classrooms. As Brown (1992) said:

> In addition to relatively standard outcome measures involving reading, writing, content knowledge, and computer competence, all of which improve significantly (Brown & Campione, 1994), this project generates a vast amount of information that is not readily subjected to standard measurement devices. We collect transcripts of children's planning, revising and teaching sessions. We collect observations of teachers' coaching and responsive teaching, as well as their direct instruction. We have records of student portfolios, including individual and group long-term projects that require them to exploit accumulating knowledge in novel ways. We score electronic mail queries to peers, teachers, and collaborators in the university community. Ethnographic observations of cooperative, or not so cooperative, interactions, such as

TABLE 1
Principles of Learning to Support Fostering a Community of Learners[a]

1. Systems and cycles
Seasonal cycles of research–share–perform activities
Supported by repetitive participant structures
Participant structures can be replaced if and only if the replacement serves the simple RSP system
Constancy at level of deep structure, variability at level of surface ritual
All activities for a purpose, a purpose for all activities
2. Metacognitive environment, reflective environment
Active, strategic nature of learning
Self-regulation and other-regulation for common good
Autocriticism, comprehension monitoring
Effort after meaning, search for understanding
Atmosphere of wondering, querying, worrying knowledge
Reflective practices
3. Discourse
Dialolgic base
Shared discourse, common knowledge
Seeding, migration, and appropriation of ideas
Mutual appropriation
An interpretive community
4. Deep content knowledge
Developmental sensitivity
Intellectually honest and demanding
Developmental corridors from children's intuitive knowledge to deep principles of a discipline of inquiry
Intermediate goals and levels of abstraction
Support, sharing
Enriched by diversity
5. Distributed expertise
Sharing for a purpose
Collaboration not just nice but necessary
Major, identity, and respect
Multiple ways in, multiple intelligences
Legitimization of differences
Community building
6. Instruction and assessment
Deliberately aligned
Based on same theory, loosely Vygotskian
Guided practice, guided participation
Multiple zones of proximal development
Transparent, authentic, and purposeful
7. Community features
Community of practice
Communities of practice with multiple overlapping roles
Link between current practice and expert practice emphasized
Elements of ownership and choice
Community beyond the classroom wall

Note. RSP = research–share–perform.
[a]From Brown and Campione (1996).

> group discussions, planning sessions, help seeking, peer tutoring and so forth are taken routinely, together with extensive video and audio taping of individuals, groups, and whole classroom settings. In fact, we have no room to store all the data, let alone time to score it. (pp. 151–152)

This wide variety of data enables Brown and Campione to study the FCL classrooms with many different lenses. As Brown (1992) pointed out:

> I find that in the interest of converging operations, and because of the multifaceted nature of my data base, I prefer a mixed approach, suiting the method to the particular data. I mix and match qualitative and quantitative methodologies in order to describe the phenomena, a mixture that is becoming commonplace in the journals, reflecting the increasingly complex issues that psychologists now address. In my own work I routinely combine a concentration on large-scale databases with in-depth microgenetic analyses of a few children or perhaps a group…. Our routine practice is to take fairly traditional pretest and posttest data from all the experimental and control students and combine that with a few selected case studies. (p. 156)

Brown (1992) worried about whether by selecting cases to illustrate her points, she may have distorted the data. She cites the example of how Bartlett (1932) misrepresented his data somewhat in his well-known studies of how people's memories for what they have read become distorted over time. As she says, "This problem of the theorist selecting those segments that prove his or her point is endemic in research that depends on transcripts or protocols culled from a large date base." (p. 162).

She also raised the issue of whether the FCL work is just the result of a Hawthorne effect. In the Hawthorne experiments the investigators found that worker productivity increased when they increased the lighting, but also when they decreased it, from which most conclude that any intervention has positive effects. She thinks that because she is showing specific learning effects one would expect to find given the intervention, that this is not a serious criticism. But she quotes the original investigators, Roethlisberger and Dickson (1939) as to why they were disappointed in their results:

> The difficulty, however, went much deeper than the personal feelings of failure of the investigators. They were entertaining two incompatible points of view. On the one hand, they were trying to maintain a controlled experiment in which they could test for the effects of *single variables* while holding all other factors constant. On the other hand, they were trying to create a human situation, which remained unaffected by their own activities. It became evident that in human situations not only was it practically impossible to keep all other factors constant, but trying to do so in itself introduced the biggest change of all; in other words, the investigators had not been

> studying an ordinary shop situation but a socially contrived situation of their own making.
>
> With this realization, the inquiry changed its character. No longer were the investigators interested in testing for the effects of single variables. In the place of a controlled experiment, they substituted the notion of a social situation, which needed to be described and understood as a system of interdependent elements. (p. 185)

This is exactly the kind of thinking that has led to the methodology of design research. Throughout their work, Brown and Campione (1996) have stressed the importance of thinking systemically about the interdependence of the elements of a design as one tries to assess its impact or modify its elements.

Diana Joseph's Passion Curriculum

The passion school concept is a design for comprehensive progressive education that uses deep learner interests to drive work on serious learning objectives. The idea was conceived initially by Roger Schank and Allan Collins as a natural synthesis of cognitive apprenticeship (Collins, Brown, & Newman, 1989) and goal-based scenario theory (Schank, Fano, Bell, & Jona, 1994). The passion school model is meant to provide specific guidance for structuring learning environments centered in classrooms organized as communities of common interests, rather than age-based communities. The defining principles for passion schools are as follows:

1. That students are assigned to curricula on the basis of their interests.
2. That students learn through active engagement in meaningful works by interacting with expert adults and more and less advanced students.
3. That through this work, learners grapple with important ideas, including adult-defined core competencies, such as those found in state and national education standards.

Diana Joseph's work takes the first step toward the development of the passion school model by shaping activity, so that learners engage in work that is meaningful to them, while at the same time grappling with important ideas. This work is being conducted through an extended, multi-phase design experiment in the context of students learning to create films as part of a "video crew." The passion curriculum design research focuses on the development of design principles, curricular structures, and strategies for activity design in the passion school model. In this section we describe the major phases in this work, to illustrate the ways in which data drives refinement and generalization in design research.

Phase 1. The first version of the video crew was carried out in an urban fifth-grade classroom from February to June 1996. The goal of the video crew was to develop videos on topics that the students were most interested in. The students worked for one hour a day three to five times a week. There were 33 students in the classroom and Joseph divided them into groups of 4 to 7 students depending on their particular interests. Their chosen video topics included fables, talk shows, sports, dance, singing, and street violence. The students in the groups were assigned different roles, with responsibilities for storyboarding, script writing, sets and props, acting, and camera operation. Because many groups had difficulties working together, one role that became critical was that of the facilitator to resolve conflicts. The curriculum provided students with guidance about their roles through structured materials and scheduled lessons with the instructor. The roles and associated guidance were designed to motivate students to work in areas of strength and weakness, depending on need as determined by the regular classroom teacher. As a culminating product, Joseph put together a video from the various pieces the different groups produced and gave copies to each of the children to take home. Joseph and several undergraduate assistants documented this process through daily field notes and videotape documentation. They interviewed participants prior to and after the Video Crew enactment.

There were several major lessons from this first attempt at the Video Crew that impacted the subsequent design. The major redesign goal was to find a means to focus the students on serious learning goals and to assess what they were learning. In order to address this issue, Joseph designed a certification system, based on the Scout merit badge system, whereby students could demonstrate their mastery of the camera, scriptwriting, acting, etc. Each certificate required students to carry out a set of activities that encompass the skills and knowledge required for mastery of video skills, as well as certain academic skills. When a certificate was acquired there were accompanying rights and responsibilities. For example, students who had earned the camera certificate had the right to take the camera out of the building for the purpose of shooting outside, but they also had the responsibility to teach others how to use the camera properly. The certification system was installed in Phase 2.

Joseph also found extensive differences across students in terms of the degree to which video production was an interest for them. This exposed a problem in the choice of setting for the study — the design research was intended to investigate how a curriculum could serve a group of students with a common interest, but video did not represent a common interest for these students. Furthermore, there was very little difference among students in terms of their experience making videos, collaborating, or creating projects. It was difficult to manage 33 students at a time, when none of the students had developed expertise, so that more expert students could help to manage the whole enterprise. Hence, it was decided to work in an afterschool setting in the next phase, so that it was possible to select students who were particularly interested in video and where the numbers were more manageable. In addition, the work called for an explicit design for training students to mentor others.

Phase 2. The next major phase of the passion curriculum design experiment recreated the video crew on a smaller scale, over a much longer period of time. The video crew prototype curriculum ran from early spring of 1997 until late spring 1998, omitting school vacations. The new design included three major categories of student activity: (1) projects which in this case produced student-made videos; (2) certifications, which linked adult-defined learning objectives with related rights and responsibilities; and (3) community life, which included student-run aspects of the classroom community. Ten students in grades four through six created a large number of complete videos, as well as a large number of plans, scripts, storyboards, treatments, raw footage, props, costumes and other video work components. Genres included original fiction, adaptation of fiction, performance, documentary, mock-umentary, animation, weather reports, and rituals. They also worked on nonvideo projects such as a Video Crew website and a number of planned social events. Their work culminated in a student-organized film festival, attended by families, friends, and teachers.

These projects were largely student-developed, though there was some teacher assignment to roles within the projects. The work was conducted with some adult guidance and support for the technical and planning aspects of video production, with content and execution strategies generally left up to the students. In addition to their project work, students worked to demonstrate expertise on certain categories of skills, through certifications. Having developed a certain level of expertise, the senior students worked to mentor novices in spring 1998. Students also participated in the community life of the Video Crew — both in terms of the traditions of video-production organizations and in terms of helping to define classroom norms. Students planned regular off-site visits to a video-editing studio, and occasional location shoots at sites such as the public library and the city zoo. In addition, they persuaded the school principal to permit weekend access to the school building for a site-specific horror-genre video.

A key area of research at this stage was student interest. Data from this phase suggested that strong student interest was unreliable. Even the most passionate students shifted in their engagement, and many students seemed to be motivated by issues other than interest in the activities. Joseph needed a more detailed theory of learner motivation in passion curricula. In collaboration with Daniel Edelson, Joseph developed a framework that used theories and findings from motivation research to organize the design of learning environments (Edelson & Joseph, in revision; Joseph & Edelson, 2002). This theory created a foundation for analysis of the Video Crew data with regard to learner motivation.

Joseph (2000) focused on field notes and artifacts to uncover patterns of learner engagement. Analytical lenses based on Edelson and Joseph's theory indicated that some students exhibited fairly consistent motivations across different activities—some were passionate about video, some pursued other interests, and some appeared to be motivated by other considerations, such as a desire to connect with friends. Artifacts produced by students with these patterns were analyzed for indications that students had grappled with ideas in the process of creating these artifacts. These findings were of interest both from the point of view of the direct work of design in the passion curriculum project, and from a more general point of view. With respect to the passion curriculum model, these findings pointed at a need for the framework to provide motivational support for students with different tendencies, through role assignment and through refinements to the certification structure. At the same time, the findings suggested the possibility that learners in general, not just in passion curricula, exhibit motivational tendencies that are fairly characteristic over time. In this sense, the passion curriculum work raises general issues, and posits frameworks, such as the Edelson and Joseph theory, that are of potential use in other contexts. This is an example of data from a design experiment raising new questions for basic research.

Phase 3. Between 2000 and 2002, the Video Crew operated in a Chicago charter school with a different cohort of students. In this afterschool setting, students chose to join video crew over other possible choices, so they were more likely to bring a strong interest in video than in the previous settings. Students attended Video Crew up to three days a week for 1–3 hours. Each day had a special focus, such as working on video projects, working on certifications and analysis of others videos. In the first year, students were trained as mentors for new students joining in the second year. Another design change was to establish the intermediate goal of showing videos at different times to other students in the school, in addition to the film festival at the end. An important innovation in this phase was to separate the roles of teacher and researcher. By working in collaboration with a teacher, Joseph separated the passion curriculum model from her own teaching practice. The core questions in this phase focused on needs for more articulated instructional strategies to complement the design strategies (Joseph, Edwards, & Harris, 2002).

The passion curriculum design framework has evolved through this design research to the point where it has a specific framework, a library of case studies describing implementation, and some evidence supporting the conjecture that the framework organizes learner motivation in the intended ways. This foundation has become the seed for the next phase of work —developing and enacting other passion curricula through other designers and teachers. A new passion curriculum, Multi-Media Design Studio, was enacted in fall of 2002, supporting further investigation of the design framework and of the motivational affordances of particular technologies (Joseph & Nacu, submitted). As an engine for the development of a learning environment, the design-research model has guided the evolution of the passion curriculum framework over time. At the same time, by demanding real-world enactment of the design, design experimentation created a setting where a key learning feature—motivation–was far more salient than it is in traditional classrooms, permitting more effective study of this phenomenon.

METHODOLOGY OF DESIGN RESEARCH

In this section we attempt to provide guidance to people who are planning to carry out design experiments. Our approach to design research requires much more effort than any one human can carry out. We put forward these ideas not because we expect each and every design experiment to embody them, but to give an overview of all the things the design-research community is responsible for. In our ideal world, design research will move in the direction of embodying many of the practices we outline here. But it will take teams of researchers and accessible archives documenting design experiments, as we discuss in the last section, to make these dreams at all possible. The guidelines we develop in this section are summarized in Table 2.

TABLE 2
Guidelines for Carrying out Design Research

Implementing a design
Identify the critical elements of the design and how they interact
Characterize how each was addressed in the implementation
Modifying a design
If elements of a design are not working, modify the design
Each modification starts a new phase
Characterize the critical elements for each phase
Describe the reasons for making the modifications
Multiple ways of analyzing the design
Cognitive
Resources
Interpersonal
Group or classroom
School or institution
Measuring dependent variables
Climate variables (e.g., engagement, cooperation, and risk taking)
Learning variables (e.g., dispositions, metacognitive, and learning strategies)
System variables (e.g., ease of adoption, sustainability, spread)
Measuring independent variables
Setting
Nature of learners
Technical support
Financial support
Professional development
Implementation path
Reporting on design research
Goals and elements of the design
Settings where implemented
Description of each phase
Outcomes found
Lessons learned
Multimedia documentation

Implementing a Design

Each implementation of an education design is different. Therefore it is important to identify the critical elements of the design and how they fit together. In order to evaluate any implementation, one needs to analyze each particular case in terms of these key elements and their interactions. Some elements will be implemented more or less as the designers intended, some will be changed to fit the circumstances, and some will not be implemented at all. What is needed is a profile for each implementation as to how each of the critical elements were implemented and how well the elements in the implementation worked together toward the designer's goals.

Brown and Campione (1996) made the argument for the FCL design that it is not important which particular activities they have designed are implemented, as long as the principles that they outline in their article are realized. Their argument implies that the principles they enunciate are the key elements of the design, and that any implementation should be evaluated as to how well these principles were implemented.

Modifying Designs as You Proceed

A goal of design research is to improve the way a design operates in practice. The teacher or researchers may see that an element of the design is not working in the course of the experiment. It is important to analyze why it is not working, and take steps to fix whatever problems appear to be the reasons for failure. In this way we collect information about failures, plus information gathered from the attempted repairs to the design, and whether they succeed or fail. It is critical to document the failures and revisions, as well as the overall results of the experiment.

The experimental methods inherited from psychology assume a fixed procedure is used throughout the experiment. Design research assumes continuous refinement. This difference has deep ramifications and requires changes in the way we analyze and report what is done. We suggest that researchers document their designs in detail, recording all major changes in design. These design changes mark the borders between phases. The goal then becomes to characterize the design elements that are in place in each phase and the reasons for the transitions from each phase to the next. Data relevant to research questions should be collected in each phase. For example, if there were 4 phases in a particular implementation, then it would be good if there were an intermediate assessment of learning outcomes between phases 2 and 3, as well as pretests and posttests. A detailed design history of this kind allows research audiences to evaluate the credibility of design decisions, and the quality of lessons learned from the research.

Multiple Ways of Looking

Rogoff (1995) called for analysts of learning environments to attend to three critical aspects: the personal layer (the experience of the individual), the interpersonal layer (one-on-one interactions), and the community layer. In the context of design experiments, researchers must additionally attend to interactions of learners with elements of the environment. There are many different aspects of what makes for an effective design, and so both designers and evaluators need to wear many hats in order to design and assess educational interventions. Consider some of the different aspects that are relevant to educational designs:

- *Cognitive level:* What do learners understand before they enter a particular learning environment, and how does that understanding change over time? Some of the tools for analysis at this level include observations of thinking through learners' representations and explanations. Through visual and verbal descriptions of ideas, researchers ask learners to expose their thinking. Are the explanations clear? Do representations capture important relationships?
- *Interpersonal level:* This viewpoint addresses how well teachers and students interact personally. Is there sharing of knowledge? Have the students bonded with each other so that they respect and help each other? Researchers use ethnographic techniques to observe these kinds of interactions.
- *Group or classroom level:* This viewpoint addresses issues of participant structure, group identity, and authority relationships. Is everyone participating? Is there a sense of the goals and identity of the group? Again, ethnography is an effective approach to analysis.
- *Resource level:* This level deals with what resources are available to learners and if they are easy to understand and use. How accessible are the resources? How well are they integrated into the activities?
- *Institutional or school level:* At this level issues arise as to communication with outside parties and support from the entire institution. Are parents happy with the design? Do administrators support it strongly? What are the micro-political issues that impact the design?

These levels are very much intertwined. To design and assess these different issues requires many different kinds of expertise: teachers, administrators, psychologists, anthropologists, media designers, etc. Conceivably one person can address all these different perspectives, but it helps to have them all represented explicitly.

Characterizing Dependent Variables

Success or failure of an innovation cannot simply be evaluated in terms of how much students learn on some criterion measure. Different kinds of evaluation are necessary for addressing questions such as: how sustainable the design is after the researchers leave, how much the design emphasizes reasoning as opposed to rote learning, how the design affects the attitudes of students, etc. To evaluate different variables, it is necessary to use a variety of evaluation techniques, including standardized pretests and posttests, survey and interview techniques, and systematic scoring of observations of the classrooms. Both qualitative and quantitative evaluations are essential parts of design-research methodology.

At least three types of dependent variables are important to assess:

1. Climate variables such as engagement, cooperation, risk taking, student control.
2. Learning variables, such as content knowledge, skills, dispositions, metacognitive strategies, learning strategies.
3. Systemic variables, such as sustainability, spread, scalability, ease of adoption, and costs.

Evaluating climate variables requires observational techniques, either by producing field notes while observing the intervention in practice, or collecting video records of the intervention and scoring those records subsequently. For example, these techniques might be used to evaluate three kinds of climate variables: the degree of engagement of students in learning in the classroom, the degree of cooperation among students in the classroom, and the degree of effort students are making to understand the curriculum topic. To evaluate these variables one might collect videos of different classes spread out over the time the teacher is carrying out the designed intervention. These videos can be scored systematically by two raters with respect to the three dimensions using a five-point scale for each 5-minute interval in the lesson. Raters would be trained using benchmark lessons for which scores have been calibrated with experts.

Learning variables are best assessed by collecting pretest and posttest measures. For example, pretests and posttests can be used to evaluate three kinds of learning variables: content, reasoning, and dispositions. To evaluate learning of content and reasoning, it is possible to use short answer or essay questions, oral interviews, or multiple-choice items. By using items from standardized tests, it is possible to compare performance to national norms for the items. To evaluate learning of dispositions, one might apply instruments developed by Dweck (1986) to assess whether there are changes in students' beliefs reflecting a move from performance goals to learning goals. There have been such changes reported in a design experiment carried out by Lamon and her colleagues (Lamon, Chan, Scardamalia, Burtis, & Brett, 1993; Scardamalia, Bereiter, and Lamon, 1994).

Systemic variables are best evaluated by interviews and surveys. For example, one might evaluate systemic variables, such as the ease of adoption of a design into the curriculum, the degree to which it is sustained in subsequent years, and the spread of use to other teachers and students. These can be measured by surveys and structured interviews with teachers and students. It is possible to develop a questionnaire that addresses advantages and difficulties teachers encountered in adopting a design in their classroom. The other variables can be evaluated by surveys administered to both teachers and students at regular intervals. The surveys will ask about what aspects of the design are being sustained and are spreading, and which aspects are not.

Characterizing Independent Variables

In evaluating any design there are a large number of independent variables that may affect the success of the design in practice. It is a matter of art to determine what aspects of the implementation situation may affect the success of the design. Our goal here is to say what general aspects of the situation researchers need to consider in order to decide what is affecting the success of the design. The contextual variables that can determine the success of an innovation include:

- *Setting.* The setting of the learning environment is a critical variable in how any design fares. The setting might vary over homes, workplaces, museums, schools, or colleges; elementary, middle or high schools; public or private schools; urban, suburban, or rural schools; elite or community colleges; and so forth. How broadly applicable an innovation is can only be determined by trying it out in many different settings.
- *Nature of the learners*. Critical variables about the learners include such things as their age, socioeconomic status, turnover rate, attendance rate, and so forth. For example, some innovations may work with weaker students and some with gifted students. So it is important to determine for which type of learners the design is effective, and in what ways.
- *Required resources and support for implementation*. In order to carry out any design, there will be a need for resources and supports of various kinds, including materials, technical support, administrative support, and parent support. If a design requires teachers to gather materials, spend time in preparation or other activities, enlist administrators or parents to make the design succeed, then these requirements need to be identified.
- *Professional development.* Often in order for a design to be successful, teachers (and perhaps others) need to be provided with professional development of various kinds. These can encompass workshops, design meetings, courses, videos of exemplary practice of the design, guided practice with expert practitioners, reflective meetings with colleagues, and so forth. Identifying what teachers need to implement the design successfully is an important aspect of designing an innovation.

- *Financial requirements*. Any intervention adds costs that need to be tracked, including equipment costs, service costs, professional support and development costs, replacement costs, and so forth. Very often substantial costs, such as technical support and replacement costs, are ignored when calculating the cost of a technological innovation.
- *Implementation path*. This term covers the variables involved in implementing a design, such as how the innovation is introduced, the time devoted to it, the duration of its usefulness, and so forth. There is a structure to the introduction and evolution of a design that needs to be characterized in analyzing any implementation.

There is a web of interrelations between independent and dependent variables. The division between the two depends on what outcomes one is interested in. But changes in any variable can have effects on other variables through complex feedback loops. Hence, changes in a dependent variable may lead to changes in an independent variable, as when increases in engagement lead to increases in attendance. The language of dependent and independent variables is only meant to capture the distinction between outcomes we should consider and those variables that may affect the outcomes.

Reporting on Design Research

The experimental literature developed a conventional structure for reporting on experiments that evolved over time. The structure consisted of four parts: Background to the problem, experimental method, results, and discussion. Because design research reconceives the experimental process, there needs to evolve a different structure for reporting. Tentatively, we propose that there should be five sections in reporting on design experiments:

- *Goals and elements of the design.* An important aspect of reporting on design experiments is to identify the critical elements of the design and how they fit together to accomplish the goals of the design. The critical elements of a design may be the materials, the activities, a set of principles, or some combination of all these. It is equally important to describe the goals of the design and how all the elements are meant to work together to attain those goals. Goals, critical elements, and their interactions need to be described in enough detail, so that it is possible to evaluate how well the design was implemented in different settings.
- *Settings where implemented.* The description of the settings needs to include all the information relevant to the success of the design outlined in the 'Characterizing independent variables' section. Differences between how the design was implemented in each setting should be detailed, so that readers can evaluate how faithfully the design was carried out in each setting.

• *Description of each phase.* The design is likely to go through a different evolution in each setting, so it is necessary to describe each phase in each setting. When changes are made in a setting, the reasons for the changes should be specified along with the effects of making the changes. It also makes sense to describe how the critical elements of the redesign accomplish the goals of the original design or how the goals have changed.

• *Outcomes found.* The outcomes should be reported in terms of a profile of values on the dependent variables in the different settings, much like qualitative and quantitative data are reported about different products in Consumer Reports. To the degree intermediate data were collected describing the different phases, these should be included.

• *Lessons learned.* Considering what happened in the different implementations, the report should attempt to pull together all the findings into a coherent picture of how the design evolved in the different settings. It is important to describe the limitations and failings of the design, as well as the successes, both in implementation and outcomes.

DESIGN RESEARCH AS SUMMATIVE EVALUATION

Although design experiments were conceived as a formative evaluation strategy, the principles involved do have implications for summative evaluation. We would argue that any assessment of educational innovations must carry out both quantitative and qualitative assessments, using comparative analysis, as does Consumer Reports. For example, if we wanted to compare how effective the Waterford and Peabody reading programs are, we would need to carry out comparative analyses in a variety of different settings, such as urban, suburban, and rural schools, and perhaps even homes, workplaces, and military settings. In such studies there must be a fixed experimental procedure, unlike the flexible design revision we recommend for formative evaluation. The assessment should produce a profile that shows the strengths and weaknesses of the designs being compared. Hence, different designs might be found to be more effective in some settings or with regard to some outcomes.

To have a sound assessment process, design-based researchers as a community should develop a consensus process to determine what variables to look at and how to assess them. The assessment should address the multiple concerns of different stakeholders, including developers, and so they should be included in the consensus process. The design-research methodology argues that we need to look at multiple contextual and dependent variables. At least three types of dependent variables are important:

1. Climate variables such as engagement, cooperation, risk taking, and student control.
2. Learning variables, such as content knowledge, skills, dispositions, metacognitive strategies, and learning strategies.
3. Systemic variables, such as sustainability, spread, scalability, ease of adoption, and costs.

As suggested previously, there are standard ways to assess each of these types of variables.

To carry out investigations fairly, the evaluators must be independent. To carry out such evaluations effectively, the country would need to invest in an independent agency, in the style of Consumers Union, with the expertise to carry out comparative evaluation. Such an agency could develop the expertise and methods for looking in a cost-effective manner at innovations in use, in a way that best informs the many different stakeholders.

CONCLUSION

Ann Brown (1992) felt that laboratory experiments, ethnographies, and large-scale studies are all valuable methodologies to study learning, but that design experiments fill a niche these methodologies do not address. It is clear from the spread of these kinds of research methods (Barab & Kirshner, 2001; Edelson, 2001; Design-based Research Collective, 2003) that design research is here to stay. But design experiments often lead to the collection of large amounts of data that go unanalyzed. Hence, it makes sense for the design-research community to establish an infrastructure that would allow researchers at other institutions to analyze the data collected in design studies, in order to address their own questions about learning and teaching. This would require the community to honor such reanalysis of data with the same status as original research and it would require research journals and tenure committees to take such work seriously. Other fields, such as child language (MacWhinney, 1995), have developed widely available archives of data, enabling researchers to discuss and analyze the same data from many different perspectives. As a community, we should strive to set up an infrastructure that can support researchers at different sites in analyzing the large data sets that design experiments are now producing.

ACKNOWLEDGMENTS

This article was written in honor of Ann Brown, whose contributions to educational research were unparalleled and whose work established design research as a viable enterprise. We particularly thank Joseph Campione for his many inputs to the description of FCL in the article, and Sasha Barab for his comments on an earlier version of the article.

REFERENCES

Aronson, E. (1978). *The jigsaw classroom.* Beverly Hills, CA: Sage.

Barab, S. A., & Kirshner, D. (Eds.) (2001) Special issue: Rethinking methodology in the learning sciences. *Journal of the Learning Sciences, 10*(1&2), 1–222.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge UK: Cambridge University Press.

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

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

Brown, A., & Campione, J. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 289–325). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brown, A., Ellery, S., & Campione, J. (1998) Creating zones of proximal development electronically. In J. G. Greeno & S. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 341–367). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Campione, J. C. (2000, April) Fostering communities of learners: Theory to practice and practice to theory. Paper presented at the American Educational Research Association annual meeting.

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

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

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Dweck, C. (1986) Motivational processes affecting leaning. *American Psychologist, 41*(10), 1040–1048.

Eckert, P. (1989) *Jocks and burnouts: Social categories and identity in high school.* New York: Teachers College Press.

Edelson, D. C. (2001) Design research: What we learn when we engage in design. *Journal of the Learning Sciences, 11*(1), 105–121.

Edelson, D. C., & Joseph, D. (in revision). Interest-driven learning: A design framework for motivating active learning. *Educational Psychologist.*

Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*(9), 27–32.

Joseph, D. (2000). *Passion as a Driver for Learning: A Framework for the Design of Interest-Centered Curricula.* Unpublished doctoral dissertation, Northwestern University, Evanston, IL.

Joseph, D., & Edelson, D. (2002). *Engineering motivation: Using research knowledge about motivation in the design of learning environments.* In P. Bell & R. Stevens (Eds.), *Proceedings of the International Conference on the Learning Sciences.* Seattle, WA.

Joseph, D., Edwards, A., & Harris, A. D. (2002, April). *Tools for building beyond the one-room progressive schoolhouse.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Joseph, D., & Nacu, D. (Submitted). Making media motivates: New media development as a motivation resource in learning environment design. *Convergence*.

Lamon, M., Chan, C., Scardamalia, M., Burtis, P. J., & Brett, C. (1993, April). *Beliefs about learning and constructive processes in reading: Effects of a computer supported intentional learning environment (CSILE).* Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

MacWhinney, B. (1995). *The Childes Project: Tools for analyzing talk* (2nd ed.). Mahwah NJ: Erlbaum.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction, 1*(2), 117–175.

Roethlisberger, F. S., & Dickson, W. J. (1939) *Management and the worker.* Cambridge, MA: Harvard University Press.

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. 252). Cambridge UK: Cambridge University Press.

Rosenbaum, J. (2001) *Beyond college for all: Career paths for the forgotten half.* New York: Russell Sage.

Scardamalia, M., Bereiter, C., & Fillion, B. (1981). *Writing for results: A sourcebook of consequential composition activities.* LaSalle, IL: Open Court.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). CSILE: Trying to bring students in world 3. In K. McGilley (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 201–228). Cambridge MA: MIT Press.

Schank, R. C., Fano, A., Bell, B., & Jona, M.(1994) The design of goal-based scenarios. *Journal of the Learning Sciences, 3*(4), 305–346.

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

Simon, H. A. (1969) *The sciences of the artificial.* Cambridge, MA: MIT Press.

Stasz, C. (2001) Assessing skills for work: two perspectives. *Oxford Economic Papers, 3,* 385–405.

Stigler, J., & Hiebert, J. (1999) *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: Free Press.