Clarifying Formative Experiments in Literacy Research

Barbara A. Bradley  
University of Kansas

David Reinking  
Clemson University

Jamie Colwell  
Old Dominion University

Leigh A. Hall  
University of North Carolina

Douglas Fisher  
San Diego State University

Nancy Frey  
San Diego State University

James F. Baumann  
University of Missouri

An increasing number of education researchers interested in closing the gap between research and practice are gravitating toward a methodological approach that conceptualizes and frames research much differently than conventional experimental or naturalistic approaches. That methodological approach, which is often referred to generically as design-based research or education design research, has been the topic of several themed issues of research journals: Educational Researcher (2003, Vol. 32, No. 1), Educational Psychologist (2004, Vol. 39, No. 4), and Journal of Learning Sciences (2004, Vol. 13, No. 1); and it has been the focus of several books (e.g., McKenney & Reeves, 2012; van den Akker, Gravemeijer, McKenney, & Nieveen, 2006), including one on its use in literacy research (Reinking & Bradley, 2008).

Further, John Easton, currently the Director of the Institute of Education Sciences (IES) in the U.S. Department of Education, has called for research that involves collaboration among practitioners, researchers, and policy makers, and that focuses on relevance and usefulness. Specifically, IES invites proposals for research that designs an intervention and then tests and redesigns it through an iterative process (Visser, 2009). Design-based research aligns well with that emphasis (Kelly, Leish, & Back, 2009; Reinking & Bradley, 2008; van den Akker et al, 2006), and IES has funded projects using that approach in the area of literacy (e.g., Baumann, Blachowicz, Maney, Graves, & Olejnik, 2009; Leu, Reinking, Hundtison, McVerry, O'Byrne, & Zawilinski, 2009). In that same vein, it has been argued that design research is an approach that would allow literacy research to make a more tangible contribution toward bettering the world (Reinking, 2011).

Literacy researchers have conducted and published research using this methodological approach, often referring to their work as formative experiments (e.g., Ivey & Broadus, 2007; Jiménez, 1997; Leski, 2001; Neuman, 1999; Reinking & Waltons, 2000). Formative experiments follow the core concepts and principles of design-based research, in some cases using a framework of guiding questions (Reinking & Bradley, 2008). Yet, many literacy researchers are unfamiliar or only vaguely familiar with this approach. Thus, one purpose of this paper is to briefly introduce formative experiments to those who know little about them and to provide several examples of how they are being used among literacy researchers, ranging from novice to more experienced. However, even among those who are using this approach in their own work, the boundaries for
conceptualizing and implementing a formative experiment in their work are not always clear. Thus, another purpose of this paper is to extend the dialog about how to conceptualize and implement formative experiments. Toward that end, after a brief introduction to formative experiments, three researchers will in turn introduce their work and will then address the following questions:

- How do you conceptualize, conduct, and report a formative experiment?
- What standards of rigor should be used to determine the validity and quality of a formative experiment?
- What knowledge, skills, or resources are needed to use this approach?
- What benefits and challenges have you encountered in using this approach?

In the final section, a senior scholar in the field with considerable experience conducting formative experiments will reflect on this approach drawing on the three examples.

FORMATIVE EXPERIMENTS

Unlike conventional approaches that have been imported into education from laboratory science or sociology, design-based research emerged among diverse researchers interested in education. Ann Brown, a prominent literacy researcher and educational psychologist, was one of the first to articulate its foundational rationales, to explore its methodological application, and to give it one of its early names: design experiments (Brown, 1992). Her development of this approach arose from her realization that the conventional experimental methods she used in her laboratory research on metacognitive aspects of reading comprehension were inadequate when trying to translate that research into viable instruction in classrooms. Dennis Newman, one of her colleagues, shared her views and conducted research in classrooms introducing the term formative experiment (Newman, 1990; see also Jacob, 1992).

As that term suggests, in this approach, an instructional intervention introduced into authentic instructional settings is modified formatively based on qualitative, and occasionally quantitative, data indicating what is or is not working and why. Another fundamental characteristic of formative experiments is that they are guided predominately by the pursuit of accomplishing a specific pedagogical goal through an intervention that can be justified as showing promise in accomplishing that goal. Although general research questions may guide data collection, the central focus is on determining how the intervention can be designed formatively to achieve a goal.

However, formative experiments, like design-based research, in general, are not only attempts to design something that works toward accomplishing a goal. Instead, they often test established pedagogical theory in the crucible of practice, and they seek to develop generalizable pedagogical theories, although these theories are what Gravemeijer and Cobb (2006) refer to as humble and local theories grounded in practice, as opposed to grand theories (e.g., theories of motivation). Ideally, formative experiments try to understand the components of an instructional intervention that are critical to success, as opposed to simply determining that one intervention works better than another or that a certain instructional move produces desirable results.

The guiding metaphors of design-based research in general, and formative experiments in particular, are engineering (designing something that works), ecology (instruction is influenced by a complex interacting web of factors), and evolution (change is slow, incremental, and established environments inherently adapt to preserve the status quo; see Reinking, 2011). Instead of asking “Which instructional intervention is best?” (the experimentalist approach) or “What is?” (the
naturalistic approach), formative experiments ask "What goal do we want to achieve, and how do we get there?" A central concept is that trying to create something that works reveals deep understandings not revealed by other approaches. Formative experiments, guided by these metaphors, respond to prominent critiques of the status of literacy research and calls for other frames and approaches (e.g., Dillon, O'Brien, & Heilman, 2000; Pressley, Graham, & Harris, 2006).

In subsequent sections, several literacy researchers relate how they have used this approach and offer their responses to the four guiding questions. The order of presentation is according to level of experience as a researcher from novice to, in the case of the concluding section, a reaction from a researcher who has been active for more than 30 years.

PROMOTING DISCIPLINARY LITERACY AND CRITICAL THINKING WITH BLOGS
JAMIE COLWELL

Description of Work

I conducted a formative experiment to investigate two pedagogical goals simultaneously: (a) to improve eighth-grade social studies students' use of disciplinary literacy skills, specifically targeting critical thinking, in a state history class; and (b) to improve preservice social studies teachers' use and understanding of instructional techniques beneficial to eighth-grade students' disciplinary literacy skills. The intervention was disciplinary-literacy instruction in both settings paired with collaborative one-on-one blogging between the middle-school students and preservice teachers to extend practice using disciplinary-literacy strategies and skills. Blogging focused on discussion about primary or secondary sources that aligned with content in the middle school state history class. I collaborated with the middle school teacher and the university instructor to plan activities related to disciplinary literacy, to integrate those activities into their instruction, and to gather data to inform how this activity might be more effectively implemented to accomplish the pedagogical goals.

How Do You Conceptualize, Conduct, and Report a Formative Experiment?

Because the literature on disciplinary literacy suggests that there are challenges to integrating it into K-12 instruction as well as into teacher education (Moje, 2008; 2010/2011; Shanahan & Shanahan, 2006), I considered a formative experiment as a useful approach to confront those challenges. Specifically, I conceptualized this formative experiment as an opportunity to simultaneously address challenges in both contexts. To me, a conceptual advantage of a formative experiment is that it permits such flexibility while simultaneously investigating how to bring about desirable outcomes in literacy instruction (Reinking & Bradley, 2008). Of particular interest in this formative experiment was setting pedagogical goals to design an intervention consistent with theory, identifying factors that enhanced or inhibited achieving the pedagogical goals, modifying the intervention to achieve the pedagogical goals, and transforming the research environments. Therefore, for this study, I selected Reinking and Bradley's (2008) framework for conducting a formative experiment. That framework also allowed for extensive collaboration with a middle-school social studies teacher and university social studies professor to determine how to integrate the intervention into the two settings without compromising non-negotiable elements of the...
Clarifying Formative Experiments

intervention (e.g., blog entries and responses to them). An important conceptual aspect of using a formative experiment was that it allowed me to focus flexibly on disciplinary literacy, while respecting the instructors’ existing curricula, teaching styles, and pedagogical philosophies. A unique conceptualization in my work is addressing two distinct, yet complementary, goals simultaneously.

What Standards of Rigor Should Be Used to Determine the Validity and Quality of a Formative Experiment?

One benchmark of rigor illustrated in my study was the selection of appropriate research sites that, at least in initial work, is conducted in classrooms where neither success nor failure is not highly likely. The university site was a four-year public university with a student population that was racially and economically diverse. The methods class consisted of typical undergraduate education majors of good, but not distinguished, academic standing. The middle-school site was located in a Title I district but was not a Title I school. Most students met or exceeded state benchmark standards for content knowledge. The class also consisted of a range of student ability levels, but no student was enrolled in special education. Thus, I believe that a thoughtful selection of an appropriate site for research is an important aspect of rigor in this approach.

What Knowledge, Skills, or Resources are Needed to Use this Approach?

One challenge I faced in my project illustrates that researchers who use this approach must be sensitive to teachers’ needs and perspectives, and must deal with them sensitively and carefully. A significant challenge was working collaboratively with the middle-school teacher to integrate disciplinary literacy instruction into her classroom. During initial planning meetings, our individual conceptualization of strategy instruction seemed compatible, and she was enthusiastic about using them. However, she did not integrate the strategies discussed in planning meetings. It was necessary to carefully document that inconsistency and perhaps determine why, but eventually it was necessary to share and discuss that conclusion with her and suggest ways of integrating more strategy instruction into her teaching. I did so by sharing my data and analysis that documented a lack of use, but in a way that showed I was trying to determine what factors might be relevant to increasing strategy use that were in her comfort zone of teaching. It worked. She understood and integrated the strategies into her curriculum the following day and into many lessons following that day. This example illustrates that researchers who conduct formative experiments need not only well-developed observational skills, but also highly developed pedagogical sensitivities and interpersonal skills.

Further, the need to develop good personal and professional relationships extended beyond the classroom. For example, it was necessary to work with individuals who managed the technical support in the district when it became clear that the site made available for blogging in the district did not allow for access outside of school.

What Benefits and Challenges Have You Encountered in Using this Approach?

One benefit of formative experiments is the collaboration with instructors in an authentic setting. In my case it was rewarding to be involved with research that allowed all stakeholders to work through the technological and instructional barriers to best determine how the intervention could work in specific circumstances. It was also rewarding to see the preservice teachers develop a better understanding of how a classroom teacher works through barriers in a project while staying in her comfort zone as a teacher. Further, it was gratifying that preservice teachers reported that they
found the intervention informative, appealing, and approachable and that disciplinary literacy was something they were more likely to consider using in their future classrooms.

However, collaboration was also a notable challenge. The type of collaboration necessary in using this approach was time-consuming, requiring frequent meetings, extensive interviewing, and observations. There was also a constant negotiation of roles during this collaboration. I found it necessary to shift from observer to participant observer to instructor in the preservice class, and these roles had to be meticulously documented to uphold standards of rigor. Even with this challenge, the intense collaboration was rewarded by improvement in disciplinary literacy skills in both populations and instructors who indicated enthusiasm to continue using similar instruction in their future classrooms, which, to me, increased the experiment's ultimate usefulness in education.

DEVELOPING EIGHTH-GRADE STUDENTS' ENGAGEMENT AND IDENTITY AS READERS
LEIGH A. HALL

Description of Work

In a yearlong study, I investigated how an eighth-grade English teacher created reading instruction for her students that accommodated their reading identities while addressing their cognitive needs. The teacher worked to achieve three pedagogical goals with the students which were to: (a) examine and alter their involvement with classroom reading practices, (b) increase their reading comprehension, and (c) allow them to progress in who they wanted to become as readers. Twenty students participated. Five students were selected and invited to participate as focal students: three whose reading comprehension was below grade level, one who was at grade level, and one who was above grade level. These focal students were selected for closer study to gain a deeper understanding of how the intervention affected students with diverse reading comprehension abilities.

How Do You Conceptualize, Conduct, and Report a Formative Experiment?

This research project was my first attempt to use a formative experiment. I decided to use this approach because I wanted flexibility to make adjustments based on the successes or challenges the students and teacher experienced. As I began to develop the goals and instruction, I realized I needed to regularly monitor how the intervention was affecting the teacher and the students. If I was truly going to work closely with the teacher and students towards achieving specific pedagogical goals, then I needed a process that was responsive to them in ways that acknowledged the realities they faced. A formative experiment provided a framework for addressing that realization.

I have found reporting a formative experiment to be a challenge. There are many stories that can be drawn from the data, and I grapple with which ones to share. However, I have realized that trying to force my report to fit into a traditional academic model of writing may not work well. Although reporting my findings has become a source of frustration, it has challenged me as a writer and scholar to consider new forms of writing and new ways to present my data.

What Standards of Rigor Should Be Used to Determine the Validity and Quality of a Formative Experiment?

In my view, the type and amount of data collected plays a significant role in determining standards of rigor and validity of a formative experiment. What kinds and how much data are
collected will vary depending on the pedagogical goals of the study. However, a critical factor of formative experiments is that they can be modified during data collection to better achieve the pedagogical goals. Therefore, a data collection plan has to create a regular assessment/feedback loop between researchers and participants around successes and challenges that are in evolvement in attempting to achieve the pedagogical goals. When changes to instruction or data collection are made, a clear record must exist that shows how the decisions were reached and why.

In my study, I engaged in classroom observations two or three times each week. The observations helped me understand how students were responding to the instruction and allowed me to document what did or did not appear to be going well in the classroom. All sessions were digitally recorded and transcribed as needed. I met separately with the classroom teacher once a week for 45-60 minutes to discuss what had or had not occurred and to plan further instruction. We reviewed the pedagogical goals, examined field notes to determine the strengths and weaknesses of the instruction, and considered if and how our instructional approach needed to be modified. I also learned about difficulties the teacher was having with the instruction and was able to use that information to make suggestions to assist him.

What Knowledge, Skills, or Resources are Needed to Use this Approach?

I believe a successful formative study requires support from both the school and the university. At the school, it is important for administrators to be supportive of the time and space teachers need to meet the demands of a study. Types of support typically needed are dedicated blocks of time to plan instruction and review data. Teachers need to know this time will be respected. Any support teachers believe they need to successfully participate should be negotiated with school administrators before agreeing to participate. Researchers can help by offering to attend meetings with administrators.

At the university level, doctoral students are a beneficial resource and can play an important role in the success of a study. For my study, they assisted with transcribing some or all of the observations, conducting student and teacher interviews, and planning sessions. They also assisted with analyzing student work and classroom interactions. It would not have been possible for me alone to keep up with transcribing these documents, because the transcriptions were needed to understand how we were successful in achieving the pedagogical goals and how we might make changes to our instruction.

What Benefits and Challenges Have You Encountered in Using this Approach?

Managing, collecting, and reviewing data required a significant amount of time, even with the support of doctoral students. I also found it challenging to step back from my study and see what was or was not working. Although being in the classroom several times a week allowed me to develop an important insider view, it also made it difficult for me to pull back and clearly assess what was happening. I asked doctoral students to periodically review recent data and offer their interpretations of what was happening in the classroom in relation to the pedagogical goals. Having fresh perspectives, somewhat more removed from the project, helped me and the teacher to gain clarity about what we needed to do.

In the end, I have a good sense of what did or did not go well for the teacher, for the students, and why. Although this type of research might seem contextualized, and not generalizable, I have
found that not to be entirely the case. Many teachers will relate to the overall issues the participating teacher faced. Additionally, I think this research stands to have an impact on classrooms, because the design allows for researchers to connect with teachers and students. It is an approach that acknowledges the human component of what we do as researchers.

CREATING DISTRICT-WIDE IMPROVEMENTS IN LITERACY
DOUGLAS FISHER AND NANCY FREY

Description of Work

Our project began when we became involved with Bonita (a pseudonym) school district. Achievement in the 44 elementary schools in the district had stalled and several schools were identified in need of program improvement, based on student achievement scores. The district leadership team and two researchers met to develop a plan to improve achievement that would affect the more than 27,000 students in the district. The approach selected was to conduct a formative experiment. Across the district, 65% of the students were Latino/Hispanic, 16% were Asian/Pacific Islanders, 14% were White, and 5% were African American. Nearly 40% of the students were current English learners, with another 32% speaking a language in addition to English at home. The agreed-upon goal was to build student competence through teacher actions. To accomplish that goal, we worked with the district leadership team to develop an instructional framework based on the gradual release of responsibility and to organize resources toward providing administrators and teachers with professional development experiences that would change teacher actions. Essentially, this formative experiment relied on teacher professional development focused on high-quality core instruction. We modified the professional development plans often, based on observations and feedback, resulting in significantly improved instruction, including a great deal of productive group work on the part of students who had been used to whole-class instruction.

How Do You Conceptualize, Conduct, and Report a Formative Experiment?

We conceptualize formative experiments as an approach that serves our purposes when we interact with school leaders who want to create lasting change. Unlike experimental or qualitative research in which the researcher asks a question, our experiences with formative experiments suggest that this approach is most effective when individuals inside school systems have a goal that they hope to reach. Goal-driven research allows for midcourse corrections in efforts to implement an intervention, which in our case is an broad plan for instructional improvement. As we conduct formative experiments, we keep extensive records about the initial efforts, what evidence was used to design the intervention, and every modification to the intervention. In doing so, we learn a great deal about what works, in which contexts, and with which students. We have reported our findings from formative experiments in several journal articles and conference sessions, but sometimes have found it necessary to present our data and findings as a case study with a traditional methodology section in line with that approach (Fisher & Frey, in press; Fisher, Frey, & Lapp, 2009). Reviewers and editors of our manuscripts, which typically unfold more like a story than a research report, have consistently asked us to report statistically all research methods in one section, despite the fact that our methods evolved during the course of the study.

What Standards of Rigor Should Be Used to Determine the Validity and Quality of a Formative
Clarifying Formative Experiments

Experiments?

In our view, documentation about each modification is critical to maintaining the credibility of a formative experiment. That documentation should include the specific reason that a modification was needed as well as the theoretical underpinnings for the next intervention. Rather than intuitive guesswork, modifications to an intervention in a formative experiment should be based on data and informed by reviews of professional literature. For example, within a formative experiment, the researcher can conduct a clinical trial, randomly assigning some participants to an intervention and others to a control condition. The data can then be used to determine the effectiveness of the intervention, modifications needed, and present an opportunity to subsequently intervene with the control group. In addition, the researcher collects information about the experiences of the individuals who have lived the experience. In other words, a phenomenological approach (Lewis & Stakes, 2010) and regular member checking (Creswell, 1997), contribute to the verisimilitude of a formative experiment. In our work, member checks have served as a way to ensure trustworthiness as we share our findings with a group of informants who discuss with us, and one another, their reflections on the findings as we present them.

What Knowledge, Skills, or Resources Are Needed to Use this Approach?

For us, the most important resource is access to an authentic environment in which we can conduct a formative experiment. That environment needs to be encapsulated by trust, as the agreed-upon intervention may change as a result of data analysis and review. In terms of knowledge and skills, the research team needs to be familiar with a wide range of methodologies as the ground regularly shifts and new tools to understand phenomena are required. In addition, we have found that formative experiments require sophisticated interpersonal skills, because situations change and trust has to be maintained. Also, all the stakeholders must commit to the pedagogical goal. Without such a commitment, the modifications to the intervention may be meaningless if different stakeholders attempt to pursue different goals.

What Benefits and Challenges Have You Encountered in Using this Approach?

In our view, the most rewarding aspect of formative experiments is the impact that researchers can have in real classrooms, schools, and districts. In our work using formative experiments, we have developed strong and lasting relationships with key stakeholders who appreciate the ongoing nature of the research and our commitment to it, rather than using their schools and classrooms as "data mines." There is a level of trust and a feeling of potential impact that occurs within a formative experiment that we have not achieved with other approaches.

In terms of challenges, we have experienced two. First, the Institutional Review Board (IRB) at our institution has been cautious about this methodological approach due to the fact that interventions can change and the research may engage in new activities and directions that have not been reviewed. That ambiguity can be addressed in the proposal to the IRB as well as in regular updates to the committee members, including subsequent requests when situations dictate change.

Second, the way we have used formative experiments requires a great deal of time. There is a long-term investment required from the outset, from developing the pedagogical goal through the multiple modifications to the intervention, all of which are important in gaining deep understanding
of what it takes to reach the goal. Nonetheless, we have found that the benefits significantly outweigh the challenges and formative experiments have resulted in new understandings and positive outcomes for schools and their students. Although it is difficult to unequivocally claim causation, key informants within the district attribute student achievement changes to the collective work we have done in meeting the pedagogical goal. They often point to the fact that 41 of the 44 schools met the criteria for being distinguished after our work and none of them remained in program improvement. Only 7 of the 44 schools met these criteria when we initiated our work with them. Grade-level proficiency in reading increased from 57% to 72% during the same period.

REFLECTIONS ON FORMATIVE EXPERIMENTS

JAMES E. BAUMANN

In this section I provide remarks about the three formative experiments just described. Specifically, I: (a) make observations about the contexts for the three formative experiments, (b) comment on the methods employed, and (c) offer a few musings on this emerging research methodology. Some of my remarks draw on information from the authors’ more elaborated conference papers.

Contexts

Cobb, Confrey, diSessa, Lehrer, and Schaeuble (2003) described five different settings, or contexts, for conducting design experiments:

- one-on-one (teacher/student) experiments;
- experiments in a teacher’s classroom guided by an external researcher;
- preservice teacher development experiments;
- inservice teacher development experiments; and
- school and district restructuring experiments.

I see three of the five contexts reflected in the studies discussed here.

Fisher and Frey’s study is a clear example of the most ambitious setting for a design experiment: the revamping of an entire school district’s elementary literacy instructional approach (Cobb et al., 2003, Setting 5). On a much smaller scale, Colwell’s study involved the collaboration of one researcher and one classroom teacher to effect changes in middle school readers’ engagement and identity, an instance of the Cobb et al. Setting 2. Colwell’s study of disciplinary literacy is a mix of the Cobb et al., Setting 3 (pre-service teacher development) and an interesting variant of Setting 2 (an intervention involving a classroom teacher, the teachers’ students, and preservice teachers). Although the majority of published formative experiments in literacy have involved the Cobb et al. (2003) second setting (e.g., Ivey & Broaddus, 2007; Reinking & Watkins, 2000), it is revealing that literacy researchers who have used formative experiments are beginning to broaden the contexts for their research.

Methods

The research methods Colwell, Hall, and Fisher and Frey employed were as varied as their research settings. Colwell used a qualitative case study design within her formative experiment. Thus, the data analyzed came from field notes, interviews, videos, and journals—the qualitative data considered to be sine qua non in formative experiments (Reinking & Bradley, 2008). Hall likewise
Clarifying Formative Experiments

gathered and analyzed qualitative data from observation field notes, but she found that the weekly meetings between her and her teacher participant to be the most fertile for acquiring insight into the intervention implementation and revision. Fisher and Frey collected the requisite qualitative data and engaged in ambitious constant-comparison analysis; however, they also collected considerable quantitative data on student achievement, and they conducted statistical analyses using that data. Thus, as formative and design experiment methods continue their evolution from obscurity to maturity across the broad educational research community (cf. Brown, 1992; Kelly, Lesh, & Biek, 2008; Sandoval & Bell, 2004; Schoenfeld, 2006; Shulman, 1997), they also are developing in our own field of literacy education (Reinking & Bradley, 2008).

Metamorphosis

In his overview of a 1997 compendium on research methods in education, Lee Shulman argued for methodological eclecticism in education research, asserting that educational researchers should "become skilled and experienced in at least two forms of research methodology" (p. 23). Shulman also argued that research questions ought to precede and guide the selection of research methodologies. As he stated,

We must avoid becoming educational researchers slavishly committed to a particular method. . . . We must first understand our problem and decide what questions we are asking, and then we must select the mode of disciplined inquiry most appropriate to those questions. (p. 24)

In her conference presentation, Hall noted that although she had a clearly articulated research question, she could not find a research design that matched that question. After considering and dismissing several methodological approaches, she stated somewhat facetiously that "formative experiments found me and became my design of last resort." Hall reported that she is now involved in a second formative experiment. We hope that formative experiments will also "find" other researchers and that it becomes their design of first choice when it suits their pedagogical research questions.

REFERENCES


Lea, D., Reinking, D., Huthison, A., McNulty, G., O’Byrne, I., & Zawilinski, L. (2009, May). Internet reciprocal teaching (IRT): A research-based model for teaching the new literacies of online reading comprehension. Symposium presented at the annual meeting of the International Reading Association, Minneapolis, MN.


Reinking, D. (2011). Beyond the laboratory and the lens: New metaphors for literacy research. *60th Yearbook of the Literacy Research Association.* Oak Creek, WI.


