USING A FORMATIVE EXPERIMENT TO STUDY HOW COMPUTERS AFFECT READING AND WRITING IN CLASSROOMS

David Reinking and J. Michael Pickle
University of Georgia

Education is not in need of research to find out how it works. It is in need of creative invention to make it work better.—Robert Ebel, past president of the American Educational Research Association, *Educational Researcher,* 1982, Vol. 11(8), p. 18.

Until recently the research investigating the use of computers in reading and writing instruction has been conducted primarily within an experimental paradigm. Following conventional approaches to educational research, researchers interested in how computer-based activities might affect reading and writing instruction have conducted relatively short-term experiments comparing treatment and control groups by gathering and analyzing quantitative data (see a review by Reinking & Bridwell-Bowles, 1991). More recently, reflecting the increasing use of nonexperimental methods in educational research (Winkler, 1992), researchers have begun to conduct long-term studies aimed at gathering qualitative data to investigate how computer-based activities affect reading and writing in classrooms (e.g., Michaels & Bruce, 1989).

Experimental studies employing quantitative methods and nonexperimental studies employing qualitative methods both have enriched our understanding about how computers might be used to benefit reading and writing instruction. Both approaches are suitable for investigating the effects of specific applications of computer technology to reading and writing instruction. However, both approaches are less suitable for determining how those applications might be implemented successfully in classrooms and schools.

Experimental studies focus on a relatively narrow range of variables. Other variables that may affect the outcome of the experiment must be controlled or presumed to produce random effects across treatments. In classroom research, it is difficult to control rigorously key variables given the dynamics of day-to-day instruction. Likewise, dismissing unanticipated events in classrooms as random events often belies their important effects on the outcome of the experiment.

Nonexperimental methods do not require control of the educational environment because they describe what *is* rather than imposing conditions to engineer what *might be.* However, findings from several long-term, qualitative studies in classrooms using computer-based instructional activities reveal another problem. Instructional outcomes tend not to be fully realized when computer-based activities are introduced into classrooms (cf. Dickinson, 1986; Mehan, 1989; Michaels & Bruce, 1989). A common

thread in these studies is that computer-based activities are not fully integrated into instruction, but remain an entity unto themselves. Moreover, rather than instigating change, innovative uses of technology are sometimes adapted to maintain existing instructional perspectives and activities. For example, Michaels & Bruce (1989) gathered quantitative and qualitative data in two sixth-grade classrooms that had been introduced to QUILL, a computer-based software package based on current views of teaching writing. After two years of ongoing data collection, the researchers concluded that "the technology did not radically reorganize the teaching and learning of writing in the classrooms. Rather the technology was shaped to fit into already established patterns of social organization and assumptions about doing and valuing writing in school" (p. 1).

In Piagetian terms, the computer-based activities were *assimilated* more than they were *accommodated* by the educational environment. Or, using the term preferred by Newman (Newman, 1990; Newman, Griffin, & Cole, 1989) the technology has not been fully *appropriated.* Computer technology is fully appropriated in classrooms only when it changes the ongoing structure of activities that existed prior to the introduction of computer-based activities. According to Newman (1990), full appropriation "may result in a new interpretation of the [computer] as well as constructive changes in classroom activity" (p. 9).

Is there an approach to research that encompasses the goals of experimental and nonexperimental studies and that simultaneously addresses how computer-based activities can be integrated into classrooms to effect positive changes in instruction? Such research is needed because there is a pervasive expectation that computer-based activities have a unique potential to restructure schooling and to promote the achievement of valued instructional goals (e.g., Cuban, 1986). In the remainder of this paper, we introduce the concept of a formative experiment (Newman, 1990), and we argue that a formative approach to research answers this question better than do conventional experimental or nonexperimental approaches. In so doing, we will draw upon our experience in conducting two year-long experiments in classrooms.

What is a Formative Experiment?

According to Newman(1990), "In a formative experiment, the researcher sets a pedagogical goal and finds out what it takes in terms of materials, organization, or changes in the technology to reach the goal" (p. 10). The pedagogical goal and the computer-based intervention aimed at accomplishing it can be derived from any existing theory of pedagogy. Because the computer-based intervention will be introduced into the complex environment of a classroom or school, it can "involve elaborate arrangements for teacher training, curriculum development, and production of classroom materials" (p. 10). However, the intervention and its plan for implementation as they are conceived at the beginning of the experiment are seen as a first draft, subject to modification during the experiment. Through systematic investigation, the researcher observes and documents factors that inhibit or enhance implementation of the intervention and the achievement of the pedagogical goal. In so doing, the researcher may gather a variety of quantitative and qualitative data.

Paralleling efforts to examine how and to what extent the pedagogical goal is being reached in a formative experiment is an investigation of how and to what extent

the educational environment is changed by the intervention. Newman states:

> Whatever the pedagogical theory motivating the experiment, the outcome to be observed must include how the environment becomes organized differently as it appropriates the technology and other resources. . . . If the environment, rather than the technology, is the unit of analysis, changes in the instructional interactions, changes in teachers' roles, and other ways that the educational environment is changed are observed. (p. 10)

Investigating change in the educational environment may entail using a variety of quantitative and qualitative methods for gathering data.

It is important to note that Newman (1990) embeds his notion of the formative experiment in a sociohistorical perspective of cognitive change derived form the Vygotskyian concept of the zone of proximal development (ZPD) (see also Newman, Griffin, & Cole, 1989). Furthermore, he sees computer technology as a means for bringing teachers' and students' interactions into the ZPD because it amplifies teachers' capability to organize the educational environment. His sociohistorical perspective also shapes his rejection of the pedagogical goal as a static endpoint for all students. He argues that many endpoints are possible and that "the environment may transcend its initial goals. It may also retain goals and organization in spite of the technology designer's concerted efforts to support alternative models" (p. 10).

Comparing a Conventional and a Formative Experiment

In this section we compare our experience in conducting two related studies in order to illustrate a formative experiment and how it addresses problems we faced in conducting a conventional experiment. Both studies investigate the effects of implementing a computer-based activity into several middle-grade classrooms. In the activity, students use a computer database to record and to share information about the books they are reading. Students are introduced to the concept of a database and are led to see how it could be used to find information about books they might enjoy reading. Then, they use the database to enter information about the books they have read and to find books that they may wish to read (Reinking & Pickle, 1992).

We view this intervention as a desirable alternative to the conventional book report—an activity inconsistent with our pedagogical theory. In our experience, required book reports may produce undesirable outcomes, especially when factors such as spelling, grammar, and the number of pages read are use to assign grades. In such cases, teachers and students assume adversative roles and students may actually read less and more narrowly (Spiegel, 1981). Nonetheless, the conventional book report is an enduring and ubiquitous activity in schools.

Over a period of years, we have been familiar with several teachers who used the computer-database activity in place of book reports. Our informal observations and discussions with these teachers indicated few of the negative outcomes associated with conventional book reports. When compared to conventional book reports, the data-base approach seemed to shift the focus from the teacher to the students, who had more of an investment in reading and writing associated with this activity. The teachers reported, and we saw positive outcomes associated with computer-based book reviews. Many students seemed to read more books and to read more diverse types of

books; they seemed to write more personal, elaborated reviews of what they had read and they seemed more concerned about the content and form of what they had written.

To research more systematically our informal observations and teachers' anecdotal reports, we designed a year-long study of three sixth-grade classrooms in the same middle school (Reinking & Pickle, 1992). Our original design proceeded from a conventional experimental paradigm in which three treatments would be compared: a condition employing conventional book reports, a condition using an off-line version of the database activity, and a condition using the computer-database activity. Because we planned to use intact classes for these treatment conditions, our design was quasi-experimental (Campbell & Stanley, 1963). We proposed to employ quantitative and qualitative methods to gather a variety of data in these three classes during the school year to determine what effects the computer database program might have on students' reading and writing when compared to each of the other treatment conditions.

In planning the project, we had a fairly clear idea of how the treatment conditions would be implemented based on our past experience, although we also realized that some flexibility would be required in designing and implementing the intervention. However, we also realized that to conduct a valid experiment we would have to make adjustments judiciously and uniformly across the three treatment groups. A major threat to the experiment's validity occurred the first week of the school year. We had arranged to work with a teacher who taught three developmental reading classes. In planning the experiment, we assumed that each developmental class would be reasonably homogeneous in regards to reading ability because there were separate remedial and gifted classes. In examining standardized test scores of students assigned to the three classes, we discovered statistically significant differences between all three classes. The teacher's perceptions verified these differences.

We did not believe that statistical control for the differences in reading ability across treatments would honestly compensate for the pervasive effect they would likely have on data collected over an entire school year. Reluctantly, we decided to implement the computer database activity in each of the three classes, to focus on possible differences by ability level between these classes, and to depend more heavily on qualitative data in reaching conclusions. Of necessity, our experiment became a one-group pretest, posttest design. Our disappointment in making this decision is reflected in Campbell and Stanley's (1963) comment that this design is "worth doing where nothing better can be done . . . [but] it is introduced here as a 'bad example' to illustrate several of the confounded extraneous variables that can jeopardize *internal* validity" (p. 7).

As the school year progressed, we found ourselves in a seemingly endless cycle of compromises that threatened the control required in a true experiment. Students did not have adequate access to books in their classroom, so we purchased more books for the classroom library. The dynamics of the classroom and the demands of the experimental tasks required that one of us become more of a teacher's aid than just a dispassionate observer. Students in each of the three classes had interesting but divergent ideas about how to enter information about the books they were reading. Several unanticipated practical problems also emerged. The daily schedule and school policies allowed students little free time to work on the computer. Class time permitted little opportunity for students to work individually on the single computer in the classroom.

Students' lack of typing skills was more of a problem than we anticipated. Further, the results we were seeing did not match our expectations. Most distressing was our conclusion that the computer did not seem to be transforming students' reading and writing as we had hoped. Unconsciously, we may have even begun to resent the teacher's reasonable needs for scheduling classroom activities when they conflicted with our needs to manage the experiment.

Each compromise seemed like a defeat in a war we were quickly losing. All the while, we were gathering and analyzing qualitative data; but, given our orientation, it seemed only to represent a meticulous recording of our experiment's demise. Gradually, we realized that we were learning something important from our difficulties. As this perspective became more lucid, we also realized that the experimental paradigm within which we approached this research project was at odds with the complexity of the educational environment. Our need to maintain control of extraneous variation was a barrier to finding and understanding the most relevant aspects of implementing the intervention and the effects it might have on the educational environment.

Our experience was leading us to conclusions consistent with the rationale for formative experiments as developed by Newman (1990) in an article that coincided with the end of our research project. Drawing on our experience and the rationale outlined by Newman, we designed and are currently involved in a formative experiment being carried out in two elementary schools. There are parallels between the completed research project and the current one. For example, the intervention is again a database book-review activity; and we are gathering similar data using similar methods. However, our orientation to research and the way we interpret the data are completely different. We are not interested in seeing if a particular intervention can have specific positive effects on reading and writing as determined by carefully controlled comparisons to alternative interventions or a control condition. Instead, we have set a pedagogical goal—increasing the amount and diversity of students' independent reading—and we are engaged in an iterative progression of activities aimed at achieving that goal. Our aim is to determine those factors that enhance or inhibit progress toward that goal. An awareness of these factors guides subsequent activities, which are continually fine tuned based on data such a surveys, observations, interviews, and focus-group discussions.

We are also looking for unanticipated effects that extend or transcend our original goal. For example, in discussions with teachers we have discovered they need additional support in working with students on the project's computer-based activities. Working with the teachers, we recruited and trained several of the students' parents to assist on the project. The increased involvement of parents has been a positive by-product of a necessary adjustment to implementing the intervention, and it is one that we did not anticipate. We are also collecting data to characterize how the educational environment is affected by the intervention. For example, we have found that some teachers are interested in extending what they are learning in the project to other aspects of their teaching, whereas others are not. Likewise, the constraints of the environment interact with the demands of implementing the intervention. For example, we have seen notable differences between classrooms with a single computer and those with access to a computer lab.

In this formative experiment, we employ several levels of experimentation simul-

taneously—what Schon (1987) has termed exploration, move-testing, and hypothesis-testing, which are all part of reflection in action. Put another way, we are simultaneously trying to determine what works, why it works, and what underlying principles might guide us in the future. For example, our planned training sessions for teachers and parents seemed to increase their anxiety when we began by showing them a refined example of what we hoped that they would be able to do after several sessions. We tested this hypothesis by looking specifically for supportive evidence and by changing our approach in subsequent training sessions.

A broad range of qualitative and quantitative data may be useful in formative experiments. Unlike conventional experimental and nonexperimental studies, ongoing data collection guides the course of a formative experiment. For example, quantitative data might suggest that students are decreasing the number of books they are reading; qualitative data may suggest why, which leads to a specific change in the way the intervention is implemented. The change is followed by more data collection, and so forth.

Advantages and Limitations of Formative Experiments

Based on our experiences, we believe a formative experiment to be more useful than a conventional experiment in studying computer-based interventions in classrooms. Formative experiments do not put the researcher at odds with the inherent complexity of classrooms and schools, as do conventional experiments. Unplanned variation and results become opportunities to learn more about implementing the intervention to accomplish pedagogical goals and about its effects on the educational environment. In short, formative experiments accommodate the flexibility and serendipity necessary for successful instruction; our conventional experiment did not.

We believe a case can be made that formative experiments are particularly suited to studying how computer-based activities might enhance literacy instruction in classrooms. First, few instructional interventions require more logistical planning than do even simple uses of computers in classrooms. Further, teachers of reading and writing, perhaps more than other teachers, may not be accustomed to using technology in their teaching. Printed and written texts are their primary instructional materials, and some may even view computers as the antithesis of the aesthetic values they wish to impart to their students (e.g, DeGroff, 1990). Thus, a substantial commitment of teachers' time and effort is required to implement pedagogically rich computer-based activities, perhaps accompanied by notable shifts in their thinking about the content and organization of instruction. The dual focus of formative experiments on identifying problems of implementation and on determining the degree to which technology is appropriated matches these conditions.

Formative experiments also break down the barriers that can separate researchers and teachers (Allen, Buchanan, Edelsky, & Norton, 1992). In conventional experiments, it is difficult not to see the teacher as a nuisance variable representing uncontrolled variation. In formative experiments, teachers' idiosyncratic responses to the intervention are viewed as useful data. Further, a formative experiment is enhanced when teachers become collaborating researchers who gather data and who help in determining what is or is not working, why, and what might be done to reach more effectively the pedagogical goal.

Formative experiments entail some limitations. They require a major commitment of time and energy on the part of researchers, who must first find schools and teachers willing to become intensely involved in a long-term project. Schools must have adequate hardware and software available to implement the proposed intervention, or the researcher must have the resources to provide what is needed. Researchers must plan and carry out methods for collecting extensive data, and they must work closely with key individuals such as principals, teachers, aides, parents, school media specialists, and students. Often, they must provide technical support and training for those individuals. As is often the case in qualitative research, it is difficult in formative experiments for the researcher to remain a detached observer.

Despite these limitations, we believe that a formative experiment provides a unique and missing perspective in researching how computers affect reading and writing in classrooms. Foremost in our support is that formative experiments address directly the need to effect improvements in classroom instruction as emphasized by Ebel in the quote that introduces this paper. Similarly, we agree with Newman, Griffin, and Cole (1989) that "the study of how educational interactions work can never be far removed from the task of engineering them to work better" (p. 147). Formative experiments create a balance between discovery and implementation in the way researchers approach their work. By engaging in such research, we are also likely to gain more credibility among those we hope will make use of our work.

REFERENCES

Allen, J., Buchanan, J., Edelsky, C., & Norton, G. (1992). Teachers as "they" at NRC: An invitation to enter the dialogue on the ethics of collaborative and non-collaborative classroom research. In C. K. Kinzer & D. J. Leu (Eds.), *Literacy research, theory, and practice: Views form many perspectives.* Forty-first yearbook of the National Reading Conference (pp. 357–365). Chicago: National Reading Conference.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Cuban, L. (1986). *Teachers and machines.* New York: Teachers College Press.

DeGroff, L. (1990). Is there a place for computers in whole language classrooms? *Reading Teacher, 43,* 568–572.

Dickinson, D. K. (1986). Cooperation, collaboration, and a computer: Integrating a computer into a first-second grade writing program. *Research in the Teaching of English, 20,* 357–378.

Ebel, R. (1982). The future of educational research. *Educational Researcher, 11*(8), 18–19.

Mehan, H. (1989). Microcomputers in classrooms: Educational technology and social practice. *Anthropology & Education Quarterly, 20,* 4–21.

Michaels, S., & Bruce, B. (1989). Classroom contexts and literacy development: How writing systems shape the teaching and learning of composition (Tech. Rep. No. 476). Urbana-Champaign, IL: Center for the Study of Reading.

Newman, D. (1990). Opportunities for research on the organizational impact of school computers. *Educational Researcher, 19*(3), 8–13.

Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive changes in school.* Cambridge, England: Cambridge University Press.

Reinking, D., & Bridwell-Bowles, L. (1991). Computers in reading and writing. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research* (Vol. 2, pp. 310–340). New York: Longman.

Reinking, D., & Pickle, J. M. (1992, May). *The effects of computer-based book reviews on middle-school students' reading and writing.* Paper presented at the meeting of the International Reading Association, Orlando, FL.

Schon, D. A. (1987). *Educating the reflective practitioner.* San Francisco: Jossey-Bass.
Spiegel, D. L. (1981). *Reading for pleasure: Guidelines.* Newark, DE: International Reading Association.

AUTHOR NOTE

This paper is related to a National Reading Research Center project of the University of Georgia and the University of Maryland (PR/AWARD NO. 117A20007) as administered by the Office of Educational Research and Improvement, U.S. Department of Education, and in part by an Elva Knight Research Award granted by the International Reading Association.

While we are grateful for their support, the findings and opinions reported do not reflect the position or policies of the National Reading Research Center, the Office of Educational Research and Improvement, the U.S. Department of Education, or the International Reading Association.