Bereiter, C. (2002). Design Research for Sustained Innovation. *Cognitive Studies, Bulletin of the Japanese Cognitive Science Society, 9*(3), 321-327.

Design Research for Sustained Innovation

Carl Bereiter

Institute for Knowledge Innovation and Technology and Ontario Institute for Studies in Education of the University of Toronto

Abstract

Although there is innovation in education it tends to be sporadic and discontinuous, with the result that innovative practices seldom win out against those with a long evolutionary history. Factors contributing to this condition include the difficulty of envisioning the human consequences of innovations and the predominance of research models that do not contribute to innovation. Design research is an emerging effort to bring what Whitehead called "disciplined progress" into education, but it has not yet taken on a clear form or purpose. Design research is not defined by its methods but by the goals of those who pursue it. Design research is constituted within communities of practice that have certain characteristics of innovativeness, responsiveness to evidence, connectivity to basic science, and dedication to continual improvement.

There is always innovation. The trick is sustained innovation, which realizes the full potential of an innovation and overcomes its original defects and limitations. Innovation, we may assume, is as old as the human species, whereas sustained innovation appears to be a product of the 19th century. It was only then, according to Alfred North Whitehead (1925/1948, p.92), that the advance of knowledge became intimately joined to the advance of technology, resulting in what he called "disciplined progress”—”a process of disciplined attack upon one difficulty after another.” The evolution of the internal combustion engine over the whole course of the 20th century and of the television receiver over the second half are both examples of sustained innovation. In both cases, the basic design was established early; there were no further dramatic innovations in design, but instead a vast number of minor innovations. The end result in both cases was a device that continued to be structurally very similar to its prototype but with enormously improved performance and reliability.

The automobile engine is especially interesting, because the inherent drawbacks of the reciprocating piston engine were recognized early and there were periodic attempts at radical innovation, none of which has yet succeeded. The Wankel engine was the most major attempt to break away from the conventional engine, and one that may yet succeed. It is a rotary engine in which the force of expanding gases is applied directly to rotating the crankshaft, instead of being applied to pushing down on a piston, whose motion must then be translated by levers and cams to rotation of the crankshaft, some of the energy of which must then be translated back into vertical motion to return the piston to its original position. Although the Wankel engine was invented in the 1920’s, the first functional engine was not produced until 1957. There followed a period of very rapid research and development, but when commercial versions began to appear in the 1970s they did not fare well. They suffered durability problems and their fuel efficiency could not be brought up to a level that made them competitive with conventional engines—this despite having Newton’s laws solidly on their side on the issue of efficiency (Hege, 2002).

The failure of the Wankel engine cannot be explained separately from the success of the conventional engine. For if the Wankel engine had been competing with the piston engine of 1910 it might very likely have won out. Instead, it was competing against the piston engine of 1970, which by then had become a very refined product. In order to be competitive, the Wankel engine would have had to undergo a similar process of refinement, perhaps taking decades. Mazda has continued development, claims to have solved the durability, emissions, and fuel economy problems, and may yet have a commercially viable product. But that will have been after 80 years of tinkering and innovation. The point is that a flawed idea, if continually refined, can win out over a better idea that has not had the benefit of as much development.

A similar story may be told in education. The role that the engine plays in an automobile is played in schools by the lesson. Lessons vary considerably depending on the subject, but their common elements are a teacher directing the simultaneous activity of a roomful of students in a sequence that moves from presentation (usually telling or showing, but sometimes discussing) to practice and application. As with the piston engine, the inherent drawbacks of the class lesson have long been recognized and there have been periodic efforts at radical innovation. The most obvious drawback of the lesson is its very limited capacity to adjust to the wide range of individual differences normally found within a school class—what detractors call its “lockstep” character. Students, regardless of ability and disposition, are required to proceed at the same rate through the same steps.

A number of schemes for individualizing instruction have arisen from time to time and gained followings, but the scheme that underwent the most extensive research and development was one called “individually prescribed instruction” (IPI). IPI relied on self-instructional materials and activities students could carry out independently or in small groups and on a system of individual study plans that allowed students to pursue different paths at different rates. IPI was intensively researched and developed at the University of Pittsburgh from about 1966 to 1972. In initial evaluations, students did not do quite as well on standardized tests as did students in conventional classes. By the end of the research program, student performance had been brought up to a level that was equal to that of students in conventional classes, but the looked-for advances beyond that level did not materialize and funding for the program dried up (Saettler, 1990). Like the Wankel engine, IPI had a compelling argument behind it. Individualization *ought* to produce better results than lockstep procedures. But, like the Wankel engine, IPI was being pitted against a technology that had undergone many more years of development—in this case, centuries more.

The discouraging lesson that might be drawn from these cases is that radical innovation has no chance in education. That is the lesson historians such as Tyack and Cuban (1997) also draw, though for different reasons. In the analysis that we offer, failures of radical innovation need not be attributed to resistances to change in the education system or in the psychological makeup of teachers. They can be attributed to the economics of innovation, which requires that an innovation pay off within a certain time frame. In the cases I have described, that time frame does not allow for the kind of sustained development that a radical innovation requires before it can be competitive.

There have, however, been radical innovations that have enjoyed sustained development, even though their pay-off was not immediate. The automobile itself is a prime example. Competing against the centuries-old technology of horse-drawn transportation, the automobile was one of the most radical innovations of all time; and it did not by any means win an instant victory. The initial automobiles were costly, noisy, and unreliable, and furthermore required an extensive infrastructure (of roads, service stations, fuel distribution networks, and so on) before they could be really practical. Nevertheless, innovation and development forged rapidly ahead. Four-wheel brakes, the fender, the electric starter, the mechanical windshield wiper, the vulcanized rubber tire—these and scores of other refinements and innovations combined to overcome the technology's initial limitations. And innovative automobile development continues, with the application of modern electronics, for instance; in 1997 Toyota and Honda, between them, patented more software than Microsoft (Aharonian, 1997).

What drove development of the automobile, as with the airplane and the digital computer, was awareness of potential. In all cases the potential was not fully grasped at the outset (much less the potential drawbacks). Indeed, the potential of computers continues to unfold as development proceeds. But enough potential could be seen in the prototypes to justify investment in further development; further development revealed further potential, and so on in a course of sustained development. This is where radical innovations in education, such as IPI, computer-assisted instruction, whole language, and online learning fail. The initial implementations, even if they generate an enthusiastic following, are not immediately successful enough to overwhelm traditional practices; but, more importantly, they do not open up a vista for further development. They do not offer the prospect of a series of innovations built on innovations, with a continually advancing horizon of possibilities.

A large part of the problem is the great difficulty of imagining human potentialities compared to imagining technical ones. It did not take great imagination for the automotive mechanic of 1900 to foresee a quiet, reliable automobile coursing along a highway at the speed of an express train. It would have taken great imagination to foresee the effects that such automobiles would have on human behavior. As we know, most predictions of the social consequences of technology turn out to be laughably wrong. But in education, the personal and social consequences of innovation *are* the potential. Unless that potential can be foreseen in some fairly clear and compelling way, there is nothing to drive sustained innovation.

Why Education Research Has Failed to Support Innovation

As Cronbach (1955) pointed out, there have been two traditions of quantitative research in education. One is correlational, and typically concerned with individual differences. The other is experimental, and typically concerned with the causal effects of "treatments" on outcomes. Neither kind of research, however, generates or lends much support to innovation. In fact, from the standpoint of educational design, the two traditions are much the same. Both involve reducing an area of inquiry to variables and then analyzing the extent to which one set of variables accounts for the variance in another set. (In the simplest case, there is only one variable in each set; e.g., does IQ correlate with scores on a creativity test? what is the effect of class size on reading achievement?) The only way to investigate design in this tradition is to treat design variations as variables. Thus, we have the classic experimental-versus-control study in which the experimental group is subjected to some innovative treatment and the control group is intended to represent normal practice. In the event that the treatments are found to produce significantly different results, the classical researcher immediately wants to know *what variables account for the difference?* Pursuing this question requires taking the design apart into variables whose separate and joint effects can be tested. Thus the direction of progress in both correlational and experimental research is toward increasingly fine separation of variables, moving farther and farther away from coherent design.

Cronbach observed, as have many observers since (Russell, 1999), that most treatment variables in educational studies are found to make little or no difference, whereas individual difference variables account for a large part of the variance in outcomes. This pervasive finding casts a pall over all efforts to design better ways of teaching. Cronbach, however, put forth the attractive idea that there may be powerful treatment effects after all, but that they are lurking in the *interaction* between individual difference and treatment variables. To simplify, treatment A may be better for some kinds of students while treatment B is better for others, with the result that, averaged out over all students, there is no discernible difference in effect. The trick, therefore, is to explore attribute-treatment interactions (ATI's) to discover the optimal matching of persons to treatments. So appealing was this idea that it sustained several decades of research and theorizing, despite hardly any success in finding powerful ATI's. Cronbach himself (1975) saw a fundamental barrier to progress in the discovery of ATI's. Suppose that treatment A and treatment B are found to have no overall difference in outcome. You hypothesize that treatment A may be better for high-achieving students and treatment B better for low-achieving ones. Suppose, however, that the hypothesized interaction fails to appear. There may be a third variable, say introversion-extraversion, which has a differential effect on how high-achieving and low-achieving students respond to treatments A and B. That is called a second-order interaction. But if that fails to appear, there could be a fourth variable, generating a still higher-level interaction, and so on. With each additional variable the number of combinations to be tested experimentally multiplies until the research becomes too vast to administer and too complex to make sense of (Cronbach vividly likened it to entering a hall of mirrors)—and yet it is still far short of coming to grips with the complexity that any discerning teacher will recognize in how different students respond to different instructional events.

Rather more successful than ATI research has been meta-analysis (Glass, McGaw, & Smith, 1981), in which a number of different studies that are judged to involve the same variable are brought together into a statistically powerful test of the effects of the variable. Educational research journals regularly carry meta-analyses on topics ranging from the effects of computer use to the effects of phonemic awareness training . Meta-analysis, however, takes quantitative research an additional step away from design relevance. In combining results from a large number of experiments in the use of educational games, for instance, all the differences among games and in ways of using them are averaged out, leaving nothing to aid the person who would like to design a more effective educational game.

To a designer, all of these approaches to variable testing have only limited interest. It is of some value to know how well treatments A and B work under various conditions, but such research tells us nothing about the relative improvability of A and B (if A represented the horseless carriage and B represented the prevailing horse-drawn carriage, B would have won, thus putting an end to that innovation). Only occasionally does comparative research provide clues as to how to produce improvements, and it can never point the way to a treatment C that would work better for everyone than treatments A or B.

In recent decades there has been movement in educational research, as in the social sciences more broadly, away from quantitative to more qualitative research, but for reasons that seldom have anything to do with advancing innovative design. The objections to quantitative classroom research have generally concerned its association with positivism—the quest for universal laws and empirical generalizations—and its aridity—its lack of contact with the lived experience of students and teachers. Although qualitative research comes in many different flavors, for present purposes we may divide it into two kinds. One kind retains the focus on variables and their interactions. It is really traditional quantitative research, minus the statistics and experimental controls, and is accordingly subject to all the criticisms expressed above. The other kind, often loosely referred to as 'ethnographic,' aims to provide an insightful account of what is actually going on. What is observed naturally depends greatly on the interests of the observer. When the observer is interested in design, much can be learned about how to improve a design by observing it in operation, and innovative designers of our acquaintance typically not only carry out such observations but do so in an ongoing collaboration with teachers to improve tools and practices. Unfortunately, however, most ethnographic researchers are naive about or uninterested in design improvement and most designers are untrained in ethnographic research, so that it is difficult to realize the potential of qualitative research for the advancement of educational innovation.

Design Research

Medical science makes use of classical correlational and experimental designs, along with meta-analyses, much the way education research does. The kind of medical advice put out through the mass media and avidly consumed by health-conscious readers is almost exclusively based on such research. It is enormously valuable. For years various herbs and dietary supplements have been purveyed with extraordinary claims for their health benefits, while doctors pooh-poohed them as worthless and possibly harmful. Little-by-little, however, medical research is sorting out valid from invalid claims, testing for side-effects, and providing a basis in evidence for the therapeutic use of certain of these traditional products. Of course, it does the same thing, and much more extensively, with new drugs and procedures. But such research bears on the *adoption* of products and procedures. Altogether different kinds of research are involved in the discovery or creation of new treatments and practices.

Although the testing of medical treatments is what makes news, the great bulk of medical and pharmaceutical research precedes the testing that occasionally produces what the evening news reports as “promising results.” It is carried out in laboratories and teaching hospitals and is typically beyond the comprehension of the lay public, which nevertheless enthusiastically supports it. What the public wants and has learned to expect is discovery and invention—new and improved cures. When governments grudgingly appropriate funds for education research, test results are all they expect.

The research that produces innovations and sustains their development has come to be called “design research.” It is any kind of research that produces findings that are fed back into further cycles of innovative design. Whitehead was referring to this kind of research when he declared, “The greatest invention of the nineteenth century was the invention of the method of invention.” As the preceding discussion suggests, design research now plays a leading role in engineering and medicine. Only recently, however, has it begun to be recognized as having a role in education (Collins, 1999). When Cronbach and Suppes (1969), under sponsorship of the U. S. National Academy of Education, produced a report on the role of research in the improvement of education, they recognized only two kinds of research: “decision-oriented research,” which deals with the making of educational choices, and “conclusion-oriented research,” which is concerned with testing hypotheses and developing theory. To this day, the idea of research-based innovation is absent from most policy documents on educatonal improvement. The assumption seems to be that innovations come from engineers and practitioners and that the practical value of research is only in producing evidence of whether the innovations are effective.

Although educational design research has been growing steadily, mainly within the learning sciences, it is still poorly understood, even within the educational research community. I heard one government official responsible for funding advanced research in education offering his explanation of design research. He said it is halfway between quantitative and qualitative research. Not only is this statement inaccurate, it is not even in the right ballpark. Design research is not defined by methodology. All sorts of methods may be employed. What defines design research is its purpose: sustained innovative development.

Successful design research requires a different kind of practice from decision-oriented and conclusion-oriented research and from the descriptive and narrative kinds of research that are becoming increasingly popular in education. To adopt the language of situated action, design research is constituted within communities of practice that have certain characteristics of innovativeness, responsiveness to evidence, connectivity to basic science, and dedication to continual improvement. The research thus constituted has distinctive characteristics such as the following:

1. Design research is carried out by or in close collaboration with designers. This is obligatory. Design research is part of the design process; if separated from it, it ceases to be design research.

2. Design research is inherently interventionist. Most educational researchers adopt a non-interventionist norm. Some, in the interests of objectivity, distance themselves as completely as possible from the educational processes they study. Others, given to “participant observation,” may interact with teachers and students, but even they are usually careful to avoid anything that could be called intervention. Design researchers,

by contrast, are trying to make something happen, and this frequently means crossing the boundary between observer and actor.

3. The most immediate goal of design research is the solution of problems formulated on the basis of perceived shortcomings and obstacles. Accordingly, design research requires a community of practice in which people both believe in what they are doing and pay close attention to negative results. This is in contrast to many educational communities that vigorously reject any negative evidence or criticism of their favored approach.

4. Design research is guided by some vision of as-yet-unrealized possibilities and is characterized by emergent goals—that is, goals that arise and evolve in the course of cycles of design and research. Design research does not go on in isolation from other kinds of research such as explanation-seeking, variable-testing, ethnographic, and advocacy research.Findings from these other kinds of research may influence the formulation of problems and goals, the generation of design ideas, and the interpretation of results. But the best design research has a visionary quality that cannot be derived from these other kinds of research, nor does it often arise from practice. It requires a research community driven by potentiality—like the engineers driven by the potentialities they perceived in the automobile, the airplane, the digital computer.

The Role of Practioners in Design Research

It goes without saying that design researchers need to work closely with practitioners. It is also obvious that the practitioners need to be ones who are receptive to innovation and willing to experiment with unproven methods. These are the requirements for a willing participant, but some practitioners can be much more than this—can in fact contribute significantly to sustained innovation.

A useful distinction has been introduced by Moore (1995). It is a distinction between two kinds of people who are receptive to innovation: “early adopters” and “visionaries.” Early adopters, according to Moore, are attracted by novelty and are eager to have the latest thing. Visionaries, as Moore describes them, are not dreamers. Rather, they are people who are attracted to an innovation because they see a way that it can help in achieving long-term goals of their own. Visionaries of this kind can spur the further development of innovations because they see potentialities that the originators of the innovation may not have seen. If you consider that most of the current uses of computers are ones that the original designers of the digital computer never imagined, you can see that there is great value in people who can see potentialities.

My own experience with innovative design research suggests that early adopters should be avoided whenever possible. They are the quickest to seize on an innovation, but they are also the quickest to abandon it in favor of the next new thing, and their approach to work with the innovation is usually superficial and unproductive. Visionaries of the kind Moore describes are rarer but well worth seeking out. They are not instantly recognizable as visionaries, because they do not trumpet their visions the way inspirational speakers do. They are people who, after you have described a design you are working on, say “I could do something with that” and then tell you what they would try to do. This can be the basis for a true partnership that carries design into new territories and to new heights.

References

Aharonian, G. (1997). 1997 US software patent statistics. Available, June 7, 2002 at http://www.derwent.com/piugl95-97/0696.html.

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

Cronbach, L. J. (1957). The two disciplines of scientific psychology. American Psychologist, 12, 671-684.

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

Cronbach, L. J., & Suppes, P. (Ed.). (1969). Research for tomorrow's schools: Disciplined inquiry for education. New York: Macmillan.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage Publications.

Hege, J. B. (2002). The Wankel rotary engine: A history. Jefferson, NC: McFarland.

Moore, G. A. (1995). Crossing the chasm: Marketing and selling high-tech products to mainstream customers. New York: Harper Business.

Russell, T. L. (1999). The no significant difference phenomenon (5th ed.). Montgomery, AL: International Distance Education Certification Center.

Saettler, P. (1990). The evolution of American educational technology. Englewood, CO: Libraries Unlimited.

Tyack, D., & Cuban, L. (1997). Tinkering toward Utopia: A century of public school reform. Cambridge, MA: Harvard University Press.

Whitehead, A. N. (1925/1948). Science and the modern world (Mentor ed.). New York: New American Library.