

Commentary

Is Research in New Technology Caught in the Same Old Trap?

Robert M. Bernard

Last week as I was basking in the leisure of the waning days of my sabbatical leave (in actual fact I was typing one of the articles that appears in this issue), a student came into my office with a question about a research design that he was analyzing. I won't go into the details of the question since it is irrelevant to the thrust of this argument. But the research question he was asking and his selection of variables brought to mind what I believe is one of the major conceptual errors that has plagued, and continues to plague, research in educational communication and technology. I will argue that the methodological contortions necessary to test the student's hypothesis are so cumbersome that the question should not be asked in the first place. Yet old lines of questioning persist, in spite of pleas from a variety of critics (Salomon & Clark, 1979; Clark, 1985; Salomon & Gardner, 1986).

The student's research problem involved comparing mean differences of achievement among three independent variables (i.e., that class of variables that are considered to be under the control of the researcher). One of the variables was gender of the student (you guessed it, the levels were male and female), a second was type of content (language content versus mathematics content) and the third was method of delivery ("computer-based instruction" versus "traditional classroom instruction")¹. The sample was comprised of male and female adolescents.

To begin with, it is questionable whether such a design could serve to exhibit the instructional potential of different delivery methods in interaction with student gender and content type. It is true that previous research has identified differential gender-related rates of skill development in language and mathematics. But it is the cause of these differences that is troublesome. If one subscribes to a biological/psychological explanation of sex differences in the two content areas (most people would not argue along such deterministic lines), a design of this type, or any instructionally-oriented design for that matter, has little hope of

¹"Traditional teaching," as used here, refers to all forms of classroom-oriented, teacher-directed instruction.

Robert M. Bernard is Editor of *CJEC* and an Associate Professor of Educational Technology at Concordia University, Montreal, Québec. I am indebted to four of my colleagues for their insightful critiques and helpful suggestions to pre-release versions of this piece. The opinions expressed herein remain, however, solely my responsibility.

doing more than simply reiterating those differences. Even for a nurture-oriented argument that implicates a complex cluster of social, attitudinal and instructional variables, this design is woefully inadequate. Only if one believes that instructional delivery makes nearly all the difference, does the research approach proposed stand a chance of demonstrating the hypothesized three-way interaction. In pointing this out, I am not criticizing the student so much as simply indicating the limited theoretical scope that his design can test. However, this is not the main point of this somewhat protracted tale.

Let's assume for a moment that we have successfully performed the mental gyrations necessary to accept the latter theoretical view -- that instruction, and particularly method of delivery, makes all the difference to achievement. What can we expect as we begin to operationalize the independent variables?

We can dispense with the content variable (language versus math) rather quickly by considering the nature of the dependent measures necessary to test differences between instructional methods in these fundamentally different areas. The proverbial "you can't compare apples with oranges" comes immediately to mind. However, just because we can't compare apples with oranges directly, doesn't mean we can't examine them descriptively. In our design, that connotes a correlational approach, not a direct comparison of means. The content areas can be dealt with, but in a fundamentally different way than was originally anticipated.

Now we come to the "horns of the dilemma" (although it may appear that we are approaching them tail first). What methodological considerations are necessary to provide a fair test of the difference between computer-based instruction (I'll call it CBI from now on) and "traditional teaching"? At first glance it seems that the two delivery methods should be comparable since a single set of objectives could be constructed to guide each. Is it true, then, that method of delivery stands as a unitary testable concept? Let's see.

Whenever human teachers are being compared with some alternative, we should immediately ask ourselves "What kind of teacher?" The answer to this question, and others, has a direct bearing on the interpretability of the the outcomes (internal validity) and how widely the results can be generalized (external validity). Since teachers come in two varieties, male and female, and since male teachers, for example, may interact differently with students of different sexes, this aspect of "teacher" should not be ignored. Sex of teacher can

Figure 1.
Counterbalancing scheme to control for student-teacher gender differences.

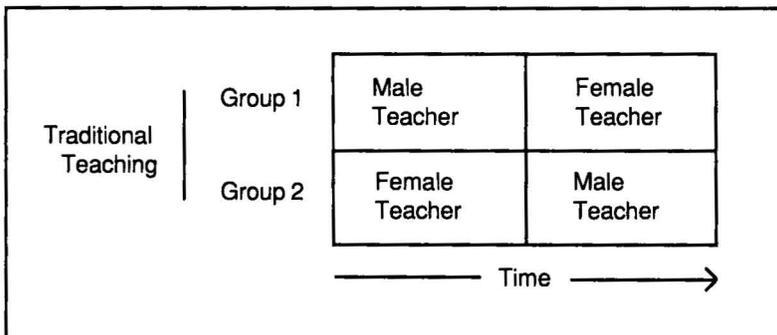


Figure 2.
2 x 2 Factorial Design in Which Student Sex and Teacher Sex are Crossed Variables.

| | | Sex of Student | |
|----------------|--------|----------------|----------------|
| | | Male | Female |
| Sex of Teacher | Male | Male Male | Female Male |
| | Female | Male Female | Male Female |

to those in the "real world"), as contrasted with the counterbalancing alternative.

So far the only criticism that can be leveled at this addition is one of increased complexity until one realizes that the gender of teacher distinction applies only to the "traditional teacher" condition (until they invent a CBI equivalent of Mr. and Ms. Pacman). This leaves what is called a partially-crossed factorial design -- crossed on "traditional teaching," but not on CBI - (Figure 2 shows the complete crossing of teacher sex and student sex, since each level of each factor is represented by a cell) or a retreat to our counterbalancing of like and opposite genders. In this alternative, differences due to gender combinations are spread over the treatments rather than isolated for measurement and analysis. This is not an unreasonable tack to take (provided we can live with the decreased external validity resulting from several teachers in the same course), but look what has happened. The differential nature of our treatments has forced us to neutralize one potentially important aspect of "traditional teaching" (teacher's gender), or face immense analytical headaches.

Consider another aspect of the teacher issue. Since we know that teachers differ from one another on many continua and that teacher effectiveness contributes somewhat to learning effectiveness, what sort of teaching characteristics should we count as important in operationalizing this aspect of "traditional teaching"? Naturally, no single teacher embodies all of the relevant characteristics of all teachers. Even if we could establish a reasonable set of criteria that defined the "ideal teacher," we would have a-devil-of-a-time finding *one*, much less one of each sex. But the real problem with our design lies not in the fact that teachers vary (many behavioral variables that are regularly researched, vary), but that teachers vary as a method of delivery, but CBI doesn't (or at least not in the same way). A similar claim might be made from the opposite direction concerning a characteristic such as length of instructional episode. CBI should proceed at a student's own pace (suggesting that length of instructional episode will vary with students), while "traditional teaching" is usually confined to a pre-set period of time with outside study time varying from individual to individual. But should student time be counted as a characteristic of method of delivery? If not, CBI varies on contact time, while "traditional teaching" does not.

The class-oriented nature of "traditional teaching" and the concomitant effects of class size (Glass & Smith, 1979; Smith & Glass, 1980) on learning is another knotty problem. Students usually work on computers independently, while "traditional teaching" is usually conducted with classes of students which may vary greatly in size. If one believes in the

be handled in one of two ways:

(a) by counterbalancing student exposure (See Figure 1) to each type of teacher, or (b) by including teacher sex as another independent variable in the design (See Figure 2). The latter approach exposes differences that may be attributable to same sex and different sex (teacher and student) combinations while preserving a modicum of external validity (external validity is considered to be high when experimental conditions are similar

socializing effects of classroom instruction (e.g., students learning from the questions and comments of other students) or the greater or lesser amounts of teacher attention that is granted by different class sizes, is it really fair to compare this human-human form of interaction with human-computer interaction?

I have touched upon but a few of the issues that a conscientious researcher would need to address in attempting to answer, unambiguously, a "which is better" question concerning our two methods of delivery. But what I have characterized here as a raft of methodological headaches (also see Clark, 1985a, 1985b), is really not that at all. It is, in my view, a not too subtle warning that two instructional treatments are so different that they shouldn't be compared in the first place (of course, two well specified and comparable methods of delivery can be compared, like two different CBI strategies). If you sensed that from the start, you might be surprised to discover that the literature of educational communication and technology is replete with comparisons of just this sort (e.g., televised teaching, programmed instruction, multi-mediated instruction). Often the finding has been "no significant differences," thank goodness. But why bother to construct, what amounts to, a unilinear ranking of instructional methods, when it is likely that each has merit under some circumstance?

Originally, I had intended to end here and, acting on that resolve, asked my colleague, Richard F. Schmid, to critique this piece. His comments are worth mentioning because I think they help to explain a few motivations that drive research of this nature. "There is, on the one hand," he said, "a natural curiosity about which of two things is better, especially when a popular view (and hope) prevails in some quarters, that one will replace the other (I am sure you know which "one" and which "other" he was referring to). On the other hand, there is a legitimate need in specific situations to know which of two (or more) instructional alternatives to select, especially when big bucks are involved."

I have little sympathy for the former view since it is engendered by the naive belief, I suspect, that a single technology (used here in the broadest sense) can ever contribute substantially to solving the "ills of instructional practice." To illustrate this view, I recently overheard a person touting the potentials of interactive videodisc for solving the "teacher problem", followed by the statement, "after all, educational media failed." In my view, educational media failed only in the minds of those who initially held unrealistic expectations for them. The use of media does solve some instructional problems, but it never could and it never will represent a general cure. The same is likely to be the case with current manifestations of instructional technology. If we fail to see them for what they are; as alternative means of achieving instructional aims, that are useful only some of the time, we are bound to be disappointed yet another time. Curiosity² is a wonderful human endowment, but it is insufficient justification, in and of itself, for attempting to answer every question that it propogates.

The latter statement -- that my previous arguments remove one means that practitioners have for selecting among instructional alternatives -- is more difficult to address, since educational research should concern itself in large measure with answering "real" questions of practice. My earlier admonitions, however, were directed towards those who believe that a general literature of comparisons among methods of delivery can service specific needs.

²In actual fact, many forces probably contribute to the pressure that is exerted for research of this sort (e.g., large institutional grants, pressure of publication, journal policies).

Here, we are no longer asking an abstract question. The conditions upon which effectiveness is largely contingent are restricted and describable. How then can a general literature ever be legitimately useful when the answer must be qualified continually with, "it depends on the specifics"?

An expensive solution to this dilemma, it seems to me, is to conduct local research that is not intended to be generalized outside of the specific circumstances of the testing site (this form of research is akin to evaluation). A far less expensive alternative lies, I think, in careful logical analysis based upon needs that are identified within a specific instructional instance. As a simple example, one would hardly choose CBI if a need for group interaction has been identified. Selection models (Romiszowski, 1976; Reiser & Gagne, 1983; Weston, 1986) represent a valuable starting point for such analysis once specific needs have been discerned. These models suggest critical features of the instructional environment, methods and materials that should be considered at each decision point.

My "solutions" are simply "off-the-cuff" answers to serious and vexing questions that plague the designers and redesigners of educational systems. In response to the critic who retorts, "decision making is not that simple," I would say, "that's true." Yet, we seem to be transfixed by the notion that research can provide "once-and-for-all", or more aptly, "one-size-fits-all" answers to these same complex instructional problems. If we are so willing to sacrifice ourselves on the alter of simplicity, we might as well go all the way, it seems to me, and simply flip a coin.

REFERENCES

- Anderson, R. H. (1976). *Selecting and developing media for instruction*. New York, NY: Van Nostrand-Reinhold.
- Clark, R. E. (1985a). Confounding in educational computing research. *Journal of Educational Computing Research*, 1, 137-148.
- Clark, R. E. (1985b). Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses. *Educational Communication and Technology Journal*, 33, (4), 249-262.
- Glass, G. V., & Smith, M. L. (1979). Meta-analysis of the research on class size and achievement. *Educational Evaluation and Policy Analysis*, 1, 2-16.
- Smith, M. L., & Glass, G. V. (1980). Meta-analysis of research on class size and its relationship to attitudes and instruction. *American Educational Research Journal*, 17, 419-433.
- Reiser, R. A., & Gagne, R. M. (1983). *Selecting media for instruction*. Englewood Cliffs, NJ: Educational Technology Publications.
- Romiszowski, A. J. (1974). *The selection and use of instructional media*. London: Kogan Page.
- Salomon, G., & Clark, R. E. Reexamining the methodology of research on media and technology in education. *Review of Educational Research*, 47, (1), 99-120.
- Salomon, G., & Gardner, H. (1986). The computer as educator: Lessons from television research. *Educational Researcher*, 15, (1), 13-19.
- Weston, C., & Cranton, P. A. (1986). Selecting instructional strategies. *Journal of Higher Education*, 57, 259-288.