Educational Technology Research That Makes a Difference Series

A Deconstructed Example of a Type 4 Study: Research to Monitor and Report on Common Uses and Shape Desired Directions

M. D. Roblyer
University of Tennessee - Chattanooga

Educational Technology Research: Addressing an Array of Challenges

The previous articles in this series of outstanding examples of educational technology studies (see Roblyer, 2005, 2006) reiterated the common plaint about the lack of useful studies and the difficulties inherent in doing meaningful, useful research in our field. Not only do we face the usual problems, obstacles, and complexities of all research on human behavior (Kaestle, 1993), educational technology also faces an array of additional challenges.

The most readily recognized roadblock to good research in our field is that of studying materials that change as they are being studied. Ours is one of the only areas of education whose tools can change dramatically in the market while we are in the midst of measuring the impact of their use. Since it takes time to plan effective methods and get approvals to do research with human subjects, this knowledge can be daunting to those who would like to make a significant contribution to the educational technology knowledge gap. Yet many researchers have managed to build research foundations for supporting significant work by focusing on concepts, strategies, and results that transcend the boundaries of specific materials. One line of research of this kind was described in the last article in this Educational Technology Research That Makes a Difference series (Roblyer, 2006). Thanks to the work of Moreno and Mayer (2002) and their colleagues, we are beginning to arrive at guidelines for developing useful multimedia products, guidelines that will be helpful regardless of the form these materials take in the future.

Another challenge to educational technology research has proven somewhat more difficult to address. In the early 1980s, researcher Richard Clark (1983) published the first trickle of what would become a torrent of criticism directed at what he called "media studies," primarily on the grounds that their underlying paradigm was flawed in ways that confounded their results. Comparing a technology-based and nontechnology-based method is pointless, he said, because the instructional method rather than the medium is what makes the difference in any study of impact. According to Bernard et al. (2004) Clark believed that "the medium is simply a neutral carrier of content and of method" (p. 381).

This criticism had a chilling effect on technology research for many years and eventually stopped the flow of evidence that technology-based strategies offer unique or important additions to effective teaching and learning. As a direct consequence of this lack of research evidence, accompanied as it was by increases in costs of technology purchases and support for school, technology advocates were confronted with a question for which they had no ready answer: Why should we spend scarce education funds on technology?

The study (Bernard et al., 2004) selected for deconstruction in this installment of the Educational Technology Research That Makes a Difference series addresses the issue of comparative effects of distance and traditional classroom learning. In their introduction, Bernard et al. cited rebuttals to Clark's criticism that may be a helpful basis for future evidence-building efforts to answer this question, thus addressing one of the most difficult and persistent challenges to educational technology research:

> Kozma argued that Clark's original assessment was based on "old, non-interactive technologies" that simply carried method and content. More recent media uses involve highly interactive set of events that occur between learner and teacher, among learners ... and even between learners and nonhuman agents or tools (p. 381).

Bernard et al. went on to give a rationale by which differences among media could be studied profitably:

> Cobb, (1997) ... argued that under certain circumstances, the efficiency of a medium ... can be judged by how much of the learner's cognitive work it performs. By this logic, some media, then, have advantages over other media, since it is "easier" to learn some things with certain media than [it is] with others ... According to this argument, the medium becomes the tool of the learner's cognitive engagement and not simply an independent and neutral means for delivering content. It is what the *learner* does with the medium that counts, not so much what the *teacher* does. These arguments suggest that media are more than just transparent, they are also transformative (p. 381). [italics added]

Based on this argument, the key to effective research methodologies for media comparison studies, then, becomes situating them in how effectively a technology-based method takes advantage of its unique capacities for interaction and engagement. The argument must be that a given instructional activity is more effective than its nontechnology alternative, because it provides a better cognitive advantage for the type of learning at hand, and because it is the technology that makes the method possible. In the Bernard et al. (2004) study, this perspective is explored and modeled to maximum advantage.

Background on Type 4 Research Studies

The introductory article to this series (Roblyer, 2005) outlined four kinds of studies that could move the educational technology field forward and that are lacking in the current published research base.

• *Type 1: Research to Establish Relative Advantage* – Studies that show that a given technology-based strategy is better than other strategies in common use because it has unique features that help bring about improved achievement, better attitudes, greater time on task, and/or more efficient learning on a topic (e.g., increasing reading comprehension through use of interactive technologies such as electronic storybooks).

• *Type 2: Research to Improve Implementation Strategies* – Studies on how to implement technology-based strategies that are already in common use so that they have greater instructional impact and benefits (e.g., implementing use of word processing for writing instruction).

• *Type 3: Research to Monitor Impact on Important Societal Goals* – Studies to find whether that technology's impact on society is positive and society-wide goals for technology are being met as originally envisioned (e.g., the goal of more equitable access to learning opportunities for underserved students).

• *Type 4: Studies That Monitor and Report on Common Uses and Shape Desired Directions* – Studies to predict and prevent negative sociological side effects of technology uses and bring about appropriate adjustments to make its overall impact on education more positive (e.g., how to address the issues and problems inherent in the current practice of students bringing handheld devices to school).

The current article explores an exemplar of a Type 4 study, though it also has aspects that provide Type 2 study results. As the introductory article in this series described, many technologies are already in such common use that what we need now is clear evidence about what sociological impact they are having on school life and whether they are meeting their own ostensible goals. This is the raison d'etre of Type 4 studies.

Certainly, no technology is more pervasive and has more potential for widespread impact than distance education. Though its use is growing rapidly and it is projected to have direct impact on the educational programs of most, if not all, instructors and students in coming years, we know relatively little about its impact on our society and how we might shape its implementation to prevent the negative society-wide outcomes that many fear are the inevitable outcome of such rapid, unplanned growth. The published study reviewed in this article offers many insights on these questions and offers some guidelines, albeit tentative, to shape implementation strategies.

A Review of a Type 4 Exemplar:
The Bernard et al. Meta-Analysis of Distance Learning Literature

Many literature reviews and meta-analyses of distance education (DE) have been published over the last 20 years. As with other reviews of technology-based applications, the consistent finding that distance environments are about equal to nondistance ones has always been accompanied by the observation that study outcomes vary widely. Some distance environments result in achievement far superior to that of classroom counterparts; some environments are demonstrably inferior. This finding cries out for explication.

Gene Glass and his colleagues introduced the technique of meta-analysis in the late 1970's as a way to provide a more statistical (and thus, interpretable) way to summarize research (Glass, McGaw, & Smith, 1981). Before then, research was summarized primarily by providing narrative descriptions or "box scores" of study results. Studies with positive results got a plus, those with negative results a minus, and those equal results were assigned an equal sign. Reviewers counted pluses, minuses, and equal signs to estimate the comparative success of the treatment. The introduction of meta-analysis offered a substantially improved estimate of impact.

Meta-analysis procedures call for calculations of effect sizes, a measure developed by Cohen (1988) to give a standardized estimate of the impact of a treatment. Effect size, known as Cohen's d, is usually calculated by subtracting the mean of the control group from that of the experimental group and dividing by the standard deviation of either group. Effect sizes are generally defined as small ($d = .2$), medium ($d = .5$), and large ($d = .8$). Not only does meta-analysis allow a statistical summary, adjusted for sample sizes of individual studies, of how much one treatment differs from another in terms of its

impact, it also (if done well) allows for a closer look at variables that may cause these differences.

Bernard et al. (2004) was published in the American Educational Research Association's *Review of Educational Research*, a journal with a long and valued tradition of providing reliable, rigorous analysis of educational trends. Bernard et al.'s meta-analysis clearly meets each of the "five pillars of high quality research," that is, criteria for good research studies referred to in the introductory article (Roblyer, 2005).

Pillar 1: The Significance Criterion

The significance criterion holds that an educational research study should make a clear and compelling case for its existence. Authors should explain why they felt the study was worth spending time and resources to pursue. Though many educators do not equate the difficulty of carrying out original research with that of reviewing it, all well-designed research – even meta-analysis – takes considerable time and resources to carry out. Thus, research should begin from the premise that a study has real potential for findings that can further the field.

In light of the fact that so many reviews of distance education research preceded theirs, Bernard, et al. take pains to make a case for why still another is needed. First, they point out that many past reviews focus on communication methods that existed at the time, e.g., mail, telephone, and television coverage. Distance education (DE) was restricted by these largely non-interactive technologies, as well as by geographical boundaries of the sources of distance education. However, the "anytime, anywhere" nature of emerging distance education has "set traditional education institutions into intense competition for the worldwide market of online learners" (p. 383). Thus, finding answers to the question of whether today's distance learning is as effective as traditional classroom learning has become even more urgent.

"Well-designed studies can suggest to administrators and policy-makers not only whether distance education is a worthwhile alterative but also in which content domains, with which learners, and under what pedagogical circumstances, and with which mix of media" (p. 383) the transformation to the distance "market" is justified by actual findings. However, after providing brief synopses of five of what they offer as the best of the previous meta-analyses, they state that they "find only fragmented and partial attempts to address the myriad of questions that might be answerable from the primary literature" (p. 386). With this as background, they conclude "it is time for a comprehensive review of the empirical literature to assess the quality of DE research literature systematically, to attempt to answer questions relating to the effectiveness of DE, and to suggest directions for future research and practice." (p. 386).

Pillar 2: The Rationale Criterion

To meet the rationale criterion, researchers should have reviewed findings from previous research for the studies they propose, and they should use these findings to generate research questions on predicted impact for their own study. It should show that the current study has a solid theory base and builds on and adds important information to past findings.

As a literature review, the Bernard et al. study uses findings from previous literature reviews, as well as critiques of those literature reviews offered by Clark, to derive a theory base for their study. It is especially significant that they credit Clark as a reviewer in their

post-article notes. Rather than circumventing or refuting Clark's criticism, they use it along with an extensive review of past findings in distance learning research (pp. 383-387) to establish the basis for their research questions and to help interpret their results. Table 1 summarizes both their research questions and why they included them.

Table 1
Bernard et al. Research Questions and Rationale for Their Inclusion in the Study

Research Questions	Rationale for Inclusion in Study
1. Overall, is interactive DE as effective, in terms of student achievement, attitudes, and retention, as its classroom-based counterparts?	Past reviews included many DE studies that focused only on non-interactive technologies, rather than the more current, interactive ones.
2. What is the nature and extent of the variability of the findings?	Past reviews found wide variation in results, ranging from very positive to very negative outcomes. Bernard et al. wanted to see if this trend continued with more recent technologies.
3. How do conditions of synchronicity and asynchronicity moderate the overall results?	Bernard et al. noted significant differences in the pedagogies possible in synchronous and asynchronous environments. Thus, they decided to see if student achievement, attitudes, and retention outcomes also differed.
4. What conditions contribute to more effective DE as compared with classroom instruction?	This kind of analysis is more helpful to practitioners than simply measuring differences between treatments.
5. To what extent do media features and pedagogical features moderate the influences of DE on student learning?	This directly addresses Clark's view that the instructional method, not the media features, accounts for differences in outcomes on criterion measures.
6. What is the methodological state of the literature?	Past reviews point out that poor research methodology hampers attempts at summarizing research results, since only studies with solid, defensible study methods can be included in a meta-analysis.
7. What are important implications for practice and future directions for research?	This question is fundamental to all research reviews, since it provides the fundamental rationale for conducting them. Results must be framed in a way that links them directly to implications for practice.

Pillar 3: The Design Criterion

The design criterion holds that the methods researchers use to study their topic must be well suited to capturing and measuring impact. Bernard et al. do an extremely thorough job of explaining and justifying their approach to meta-analysis. A meta-analysis is only as good as the methods it uses to select studies for inclusion; methods must be logical, comprehensive, and rigorously applied. Bernard et al. provide a detailed description of the keyword combinations and sources they used to search for studies. They also stated extensive criteria for including a given study, criteria that must have taken considerable time to ascertain for each study being considered. To be included, a study had to:

- Involve an empirical comparison – Studies comparing DE with national standards or norms, rather than using a control condition, were excluded.
- Have DE as a primary condition – Studies in which instruction was done with face-to-face meetings more than 50% of the time were excluded.
- Report outcomes for both experimental and control groups – Sufficient data had to be available to calculate effect sizes.
- Be publicly available or archived – This makes their methods and results more replicable.
- Include at least one achievement, attitude, or retention outcome measure
- Specify the type of learner involved
- Be published between 1985 and 2002 – This was the only criterion that could have been narrowed somewhat, since Internet-based distance methods only came into common use after about 1995.
- Include outcome measures that were the same or comparable – Both experimental and control groups had to have used the same exam or other outcome measure.
- Include outcome measures for individual courses, rather than entire programs – This allowed for better scrutiny of study variables.
- Include only the published source when data about a study were available from different sources – This was apparently designed to assure results were from primary sources, rather than reported second- or third-hand (p. 389).

One final notable procedure was the omission of outliers, a common characteristic of well-designed meta-analyses. Although studies that achieve unusually higher or lower results than most other studies are, perhaps, worthy of separate qualitative analysis, they are assumed to be not typical of usual methods. Omitting them helps make the remaining results more statistically "combinable."

Pillar 4: The Comprehensive Reporting Criterion

The comprehensive reporting criterion says that a research article must offer sufficiently detailed information to allow others to analyze and build on previous work. On this criterion, arguably the most important of the five for a meta-analysis, the Bernard et al. team excelled. Not only do they do a thorough job of describing their procedures, in an eight-page results section, followed by a 10-page discussion of results, they explained their findings in meticulous detail and with compelling clarity.

They reported findings by synchronous and asynchronous studies, and summarized these two groups by the three criterion measures: achievement, student attitudes, and retention. They also reported results in tables of verbal description, tables of numerical information, and graphic (distributions of effects sizes) formats, thus contributing even more to easy reading and digesting of information. Some of their findings that qualify as remarkable are summarized in Table 2.

The finding of substantially better results for asynchronous DE is most interesting and bears including as a variable in future reviews of this kind. Most current DE tends to use Internet-based course management systems, which support primarily asynchronous methods. However, some locations are experimenting with IP-based videoconferencing as a "distance strategy of the future." Since videoconferencing is designed to support synchronous environments, which Bernard et al. found consistently yielded lower effects, practitioners may want to take note of these results and consider using it as a course supplement for classroom or asynchronous DE, rather than a primary course delivery system.

Table 2
Summary of Selected Bernard et al. Study Results

Category of Results	Findings

Overall: By course characteristics	• DE effects were large (i.e., DE was better than classroom instruction):
	    ○ When efficient delivery/cost savings was the reason for offering it
	    ○ For K-12 students
	    ○ In military and business subject matters courses
	• Math, science, and engineering subjects worked better in the (FTF) classroom
	• Computing and military/business courses worked better in DE

Overall: By research methodology	• Wide variability in methods and outcomes • Methodological weaknesses in studies were prevalent
Overall: By DE and other methodology	• Larger effects in asynchronous (as compared with synchronous) DE • Pedagogical methods account for most differences between DE and classroom results • Active learning (problem-based and with collaboration among students) fosters better achievement and attitudes, but only in asynchronous DE
Achievement	• Consistently favors DE in asynchronous environments only
Attitude	• Consistently favors classroom instruction • Substantially better effects in asynchronous (as compared with synchronous) DE
Retention	• Consistently favors classroom instruction • Substantially better effects in asynchronous (as compared with synchronous) DE

It may also prove unsettling to DE proponents that both attitude and retention results usually favor non-DE environments. But certainly the biggest bombshell in Bernard et al. findings was the support for Clark's long-held observation that instructional method, rather than any distance technology feature, was the most important single contributor to criterion outcomes. They emphasized the instructional design adage that a medium (e.g., distance learning) should be selected in light of an instructional practice that requires it, and not the other way around. In light of the current rush to be competitive in the distance market, this may be difficult advice for organizations to heed.

Pillar 5: The Cumulativity Criterion

A literature review and meta-analysis is a cumulative activity by definition. While it is not clear whether or not they plan to do further meta-analyses as more good studies of sound pedagogies using interactive technologies become available, Bernard et al. do offer some helpful advice to other would-be reviewers: "continuing to compare DE with the classroom without attempting to answers the attendant concerns of 'why' and 'under what conditions' represents wasted time and effort" (p. 416).

Invitation to Nominate Exemplary Studies

The article reported here contributes in important ways to the research foundation that is key to making the case for funding and use of educational technology. It offers benefits of several kinds. First, it models sound practice for research reviews, especially those that use meta-analysis, using a clear and well-articulated theoretical and research foundation and defensible, replicable methods. Second, it models the kind of clear, detailed reporting of methods and results that shows that its findings are as valid and reliable as possible. Finally, it offers sound, if still somewhat tentative, guidelines for practice in an important, expanding area of technology: distance education. These guidelines are especially noteworthy since they address longstanding criticism by Clark and others in a way that allows useful comparisons of technology-based (i.e., DE) and non-technology based (i.e., classroom or non-DE) applications.

As did the introductory articles in this series, this article ends with an invitation to all educators in the field of educational technology and in the content areas to nominate studies of similar high quality to serve as exemplars of the criteria described here. We would like to include examples of the other two of the four types of studies as reflected in content-area research. Nominations may be submitted to CITE editors for inclusion in this series.

References

Bernard, R., Abrami, P., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P., Fiset, M., & Huang, B. (2004). How does distance learning compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379-434.

Clark, R.E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445-459.

Cobb, T. (1997). Cognitive efficiency: Toward a revised theory of media. *Educational Technology Research and Development, 45*(4), 21-35.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Kaestle, C. (1993). The awful reputation of educational research. *Educational Researcher, 22*(1), 23-30.

Moreno, R., & Mayer, R. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94*(1), 156–163.

Roblyer, M. D. (2005). Educational technology research that makes a difference: Series introduction. *Contemporary Issues in Technology and Teacher Education* [Online serial], *5*(2). Retrieved February 1, 2007 from http://www.citejournal.org/vol5/iss2/seminal/article1.cfm

Roblyer, M. D. (2006). A deconstructed example of a type 2 study: Research to improve implementation strategies. *Contemporary Issues in Technology and Teacher Education* [Online serial], *6*(3). Retrieved February 1, 2007, from http://www.citejournal.org/vol6/iss3/seminal/article1.cfm

Author Note:

M. D. Roblyer
University of Tennessee - Chattanooga
email: margaret-roblyer@utc.edu