

What Does Research Mean to You?

Making Educational Technology Research Relevant to Educators

The U.S. government's No Child Left Behind (NCLB, www.nclb.gov) legislation emphasizes the use of scientifically validated learning resources. This requirement has stimulated much discussion about how to define educational research. For the past two years, the Center for Applied Research in Educational Technology (CARET, <http://>

By John Cradler, Ruthmary Cradler, and Robert Clarke

Subject: Research methodology, NCLB, accountability

Audience: Researchers, administrators, technology coordinators, teacher educators, teachers

Grade Level: K–12 (Ages 5–18)

Standards: *NETS•T* III, V; *NETS•A* I, III, IV (www.iste.org/standards)

caret.iste.org) has screened and analyzed more than 700 studies to determine for each the scientific rigor (methodology and statistics) of the research and its applicability to instructional practice.

The CARET review process can be found on the CARET Web site and in previous Research Windows articles in *L&L*. As a result of the CARET reviews, as well as the experience of the authors, some specific issues have emerged that are relevant to both the education community and developers of electronic learning resources intended for use in K–12 education. Similar issues were raised in the 1960s and 1970s, but the concerns expressed then with regard to scientific rigor and quantitative methods subsequently diminished until the authorization of NCLB. A discussion of several of these issues and some suggestions for addressing them are the main focus of this article.

Statistical Significance and Educational Significance

For this discussion, the expression *statistical significance* refers only to statistical conclusions that are based on the analysis of data obtained during the course of the research. It has to do with how sure one can be that something really happened, that the observed statistical results a difference in mean test scores between two groups (e.g., occurred because of more than chance alone). The term *educational significance* is another matter. Given that the research methodology is adequate and the results are statistically significant, how relevant or important are the results for educational practice, either in general or in a specific context?

Much of the research available and examined by CARET is descriptive and based on surveys, interviews, and ethnographies and case studies. Although



the educational significance of what was discovered in many of these studies is emphasized, the research methodology and statistical treatment falls short of what is demanded by NCLB. On the other hand, there are methodologically rigorous studies in which the use of appropriate statistical methods demonstrates that large groups of students using computers or video significantly, in the statistical sense, outperformed their randomly selected control-group counterparts. The difficulty with many of the statistical studies is that they do not provide a sufficient basis for consumers (e.g., school districts) to evaluate the educational relevance of the results. That is, would a local replication of the essential aspects of

the studied technology be feasible in one's own district and have a sufficiently significant, in the educational sense, effect on student performance in target areas to justify the time, effort, and expense involved?

Statistical significance, as mentioned above, refers to the statistical conclusion that the results cannot be accounted for by chance alone, and thus it is reasonable to conclude that the intervention being studied (e.g., computer-assisted learning) may well have had an effect. Statistical significance is not a direct measure of the size of that effect. Unfortunately, an all-too-common misinterpretation is that statistical significance implies effects large enough to be important. There are a variety of reasons why this is a misinterpretation, the most important of which concerns the size of the samples used in the study (Slavin, 2003). Because of the influence of sample size, true effects of the intervention, too small to be of any practical importance, may result in statistical significance when very large samples are used. Conversely, effects that are large enough to be of practical importance, may not result in statistical significance when small samples are used.

The forgoing is not an argument against the use of large samples, for larger samples can give a more accurate picture of the true effect of the intervention. Nor is it an argument against the use of significance testing, for a significance test provides an invaluable tool for preventing us from taking too seriously those observed sample effects that could well be from chance alone—a sort of statistical screening device. Rather, it is an argument for looking beyond the outcomes of any significance tests that have been performed to see what the sample results can tell us about the likely sizes of the true effects and to assess the educational significance of those effects.

During the 1960s and 1970s, the U.S. Department of Education's Joint Dissemination Review Panel (JDRP),

in a search for programs with convincing research on their effectiveness, found that statistical significance alone was not convincing enough and added the concept of educational significance to the program and project review criteria. They found that “educational significance” documentation provided the information about the educational implications and relevance of the intervention that is needed to inform decisions about how to replicate the intervention. (Tallmadge, 1977). The JDRP concluded that a balance was needed between methodological and statistical rigor on one hand and educational significance on the other. The CARET review criteria include criteria pertaining to both research design and statistics as well as educational significance. This made it possible for CARET staff to profile studies and classify them according to the degree and nature of their potential education implications.

Enhancing the Educational Significance of a Research Study

There are a number of ways an investigator can increase the educational significance of a research study and in so doing, assist consumers in assessing the relevance of what was found for their own local situations. Among these are the following:

1. *Include tests that directly assess the skills or knowledge intended to be affected by the technology being studied.* Studies that use only standardized tests may not accurately measure actual student change. Standardized achievement tests are designed to be curriculum-neutral and may not directly relate to specific applications of a program or technology application, thus making it difficult to formulate specific instructional implications of the study. Put in terms of medicine, “Standardized testing may provide the vital signs, but deeper diagnostics are required if the patient requires further interventions” (Serim, 2002, p. 6).

2. *Identify and consider the cost of the technology application in terms of teacher, administrator, and support staff time; training; hardware and software installation and maintenance; connectivity; and technical support.* Major factors of educational significance are the costs of replicating locally the technology application or program examined by the research. Such costs include not only the obvious costs of hardware, software, and getting the equipment up and running, but also the less obvious costs of all the other support resources needed in a school setting. These costs clearly have implications for educators. Unfortunately, all the factors, including all the support necessary for valid replication in a local situation may be neither considered nor addressed in technology applications examined in controlled research situations, and thus their costs are easily overlooked.
3. *Identify and describe in detail the elements of the specific technology application being studied.* If possible, use a research design and employ a statistical analysis that will shed light on which elements make the greatest contributions to instructional outcomes. This will help others ascertain which program elements are most important to consider replicating locally to obtain maximum instructional benefit. In addition to equipment and specific instructional design features that are built into the application, program elements may include optimal length of time and frequency of use of specific program features. Other features might be options or extensions of the application, such as print materials to accompany the software, closed caption options on video, and resources for homework or after-school programs.
4. *Specify the overall context within which the technology application has been or is being implemented and studied.* Research consistently shows that technology, like any other instructional tool or strategy, is effective to the extent that basic conditions are clearly established for teaching and learning to occur. Among those that are important to describe are the degrees to which:
 - Instructional planning incorporates technology
 - There is leadership and support for the use of technology
 - Staff development is directly relevant to the application
 - There is access to appropriate hardware and software
 - The technology application supports or extends other curricular programs and priorities
5. *Describe in sufficient detail any other contextual or implementation variables that are likely to be related to student outcomes.* CARET finds that the more information about the intervention that is provided, the more that specific implications and suggestions can be drawn and applied to teaching practice. Educators find studies that provide practical implications to be of the greatest use. For instance, if a study simply reports that a particular technology application results in significant gains in achievement, it is difficult for a potential consumer to make an adequately informed decision concerning a possible application in his or her own situation. The educational significance is enhanced when the research report describes the type of students for whom the program is best suited, the optimal time to use it, the desired core instruction for which the program is a supplement, the prerequisite student skills needed to use the technology, and so on. The importance of each of these implementation variables in relation to the statistically significant outcomes is a necessary condition for being able to replicate the program or intervention strategy.
6. *During the planning process, use of a visual approach can help clarify possible cause and effect relationships inherent in the technology application.* The possible roles played by various aspects of a proposed program can sometimes be clarified by using a flowchart-like representation known as a causal map. (*Editor's note:* Read more about causal maps in "What Is Causal Mapping," *L&L* 29[7], p. 9.) In this way, the potential cause and effect relationships among the variables, as well as their possible relationships to instructional outcomes, can be represented visually. Such a visual representation encourages conceiving of the program as a system and can facilitate the design of an educationally significant study. With this process, the investigators describe conditions or context of the educational system as well as any preparation such as professional development of the teachers and prerequisite skills needed by students. All aspects of the intervention or technology application are described, as are anticipated outcomes. Then a causal map can be drawn that illustrates the potential cause and effect relationships. When the context is clear and all aspects of the implementation are defined, intermediate assessment tools can be developed to collect the implementation data needed for assessing the relationship of these variables to the outcome or summative evaluation data. This process was tested on a large-scale Technology Innovation Challenge Grant project (Just in Time) in Idaho. The process provided a structure for teachers to more clearly define the factors leading to the student outcomes and thereby identifying strategies that could be applied to their teaching practice. As a result, there was a high positive correlation between quality

of the teacher's causal maps and the resulting student projects (Cradler, 2002). Figure 1 is a simple example of a causal map showing the relationship between formative and summative evaluation. The educational significance of a study is most reliable when it is based on what is known as formative evaluation or sometimes process evaluation. Formative evaluation looks at the process and defines the specific instructional activities needed to replicate the program or intervention under study.

Importance of Background Research

Research and development in education should build on what has been developed and tested in the past. Unfortunately, many studies or evaluations of technology applications did not provide adequate coverage of background research that may have—and ideally should have—provided guidance for the design and execution of the study. In some of these instances, the investigators may have been unfamiliar with any relevant background research. Such cases are very unfortunate, because prior research can provide valuable information concerning the most unique and effective contributions technology can potentially make in the instruc-

tional area under study—what has worked and what hasn't worked and in what kinds of situations—as well as information concerning the pedagogy to be enhanced, research methodology and statistics, assessment methods, and so forth. Evaluations of technology implementation projects, such as federal-, state-, or foundation-supported grants can be a rich source of lessons learned about program implementation and initial indicators of change in teaching and learning. Even case studies of limited size can be used to provide clues to areas for more extensive experimental research or triangulated research with multiple methods of data collection and analysis. This approach is common in medical and other research and should be applied to studies on educational technology. Familiarity with such information on the part of the investigator greatly increases the likelihood that he or she can design and carry out a study that will be both statistically and educationally significant. It is important that in any reporting, the investigator include an adequate account of prior work that may have informed the study. This will facilitate understanding and interpretation of the study and assist future investigators concerned with the same or similar applications of technology.

Asking the Right Questions about Technology

CARET reviewers often find that studies of technology simply ask whether the presence of computers and/or Internet access improved student learning. These studies then suggest that technology, in some general way, was responsible for a change when, in fact, it was a specific application of technology that made the difference. Conversely, studies can suggest that technology did not make a difference when, in fact, it did when specific and appropriate uses were singled out from the overall data. The National Research Council has identified characteristics for scientific research. Among the recommendations are to pose significant questions that are sufficiently specific and answerable using appropriate objective methodological and statistical tools (Feuer & Towne, 2002). The questions asked and investigated should be significant or relevant and directly linked to the objectives of the program being evaluated, and the study should incorporate a design that documents both statistical significance and the educational significance of the program or technology application. CARET has found that most studies on educational technology answered questions about context and conditions under which the technology was used rather than its intended educational use. Asking the right questions helps the investigator decide what other factors or conditions need to be controlled and/or considered when designing the study.

The Bottom Line

To the extent that relevant studies address the sorts of critical issues described above, there is no question that research can provide a valid assessment of the possible contributions of various technologies to teaching and learning. However, a study on the effects of technology is meaningful and useful only to the degree that specific instructional

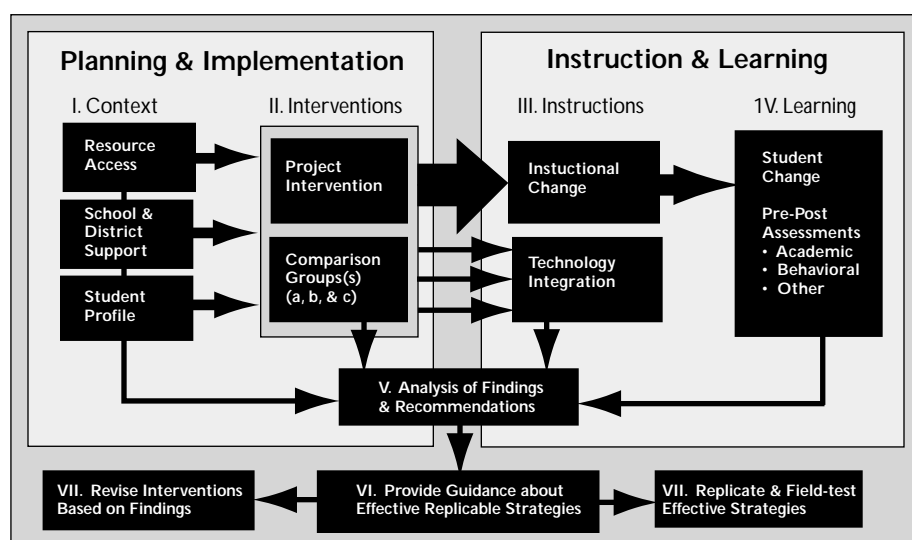


Figure 1. A causal map of the relationships between formative and summative evaluation.

Research continued on page 59.

Research continued from page 53.

goals have been clearly defined and that the technology application is directed toward those goals. Studies on the effects of technology provide the most reliable and valid results when both the statistical and educational significance of the application or intervention are relevant to the specific use of the technology being studied. The results of the study must then be referred back to the specific aspects of the application and the conditions under which it was implemented. Broad statements such as: "the research showed that the presence of computers resulted in a statistically significant improvement in student learning" are, without the necessary specifics, meaningless and possibly very misleading. There is no completely accurate and objective statistical way of generalizing the effects of a particular technology such as computers beyond the specifics of the situation in which it was tested (e.g., instructional goals, types of students, nature of the particular application under consideration). Generalizing from the research study to the local situation is ultimately a matter of professional judgment, the validity of which depends on knowledge of the specifics of both situations as well as an understanding of the sorts of issues described above.

Some Conclusions about Research and Effects of Technology

It has been shown that the addition of technology as an instructional tool and resource can make significant contributions to teaching and learning to the extent that the study addresses the issues discussed in this article. Studies on the effects of technology provide the most reliable and valid results when both the statistical and educational significance of the intervention or technology application are relevant to the specific use of the technology being studied. The results must be related back to the specific aspects of the intervention and the conditions under which it was implemented. The effects of technology can-

not be generalized and can only be described in terms of a specific application studied under defined conditions. Additionally, the research methods need to be appropriate to the carefully framed questions being addressed, which means that studies will use a combination of quantitative methods and description, depending on the questions intended to be answered by the study.

Therefore, it is very important that educators be able to analyze and evaluate research accordingly before drawing conclusions about the implications of the results for use in local learning environments. This becomes especially important now that the U.S. government is applying a great deal of pressure to the educational and business communities to show statistically significant effects of technology-based interventions on teaching and learning. It becomes easy for the consumer and the investigator to become overly impressed with statistical results and forget that statistics provide only part of the answer.

Recommendations and Suggestions

With current emphasis on documenting the effects of educational interventions and technology applications on teaching and learning, we make the following recommendations for researchers, developers, and educators:

1. To create common understanding among the research, developer, and K-12 communities, representatives from all three should be involved in the design and development of strategies for conducting and analyzing research and disseminating findings.
2. Thorough reviews by researchers and evaluators of relevant studies or reports from the research, evaluation, and practitioner communities can greatly assist the design of research to gain greater educational relevance.
3. Educators and publishers need to understand the importance of conducting studies that are both statistically and educationally significant.

4. Practicing educators need to be involved in the review and analysis of research on technology to increase the extent to which findings can be generalized and applied.

References

- Cradler, R. M. (2002). *Just in Time Challenge Grant: Final evaluation report 1997-2002*. San Mateo, CA: Educational Support Systems.
- Feuer, M., & Towne, L. (2002). *The logic and basic principles of scientific based research*. Washington, DC: National Research Council. Available: www.ed.gov/offices/OESE/esea/research/feuer-towne-paper.html.
- Serim, F. (2002). *Strong medicine: Scientifically based research and school practice*. Washington, DC: Consortium for School Networking. Available: www.cosn.org/members/compendium_issue_one/article_1.pdf.
- Slavin, R. (2003). A reader's guide to scientifically based research. *Educational Leadership*, 60(5), 12-16.
- Tallmadge, G. K. (1977). *The Joint Dissemination Review Panel ideabook*. Washington, DC: U.S. Government Printing Office.



John Cradler is the co-director of the CARET project and president of Educational Support Systems, Inc. (ESS). He first became involved with technology when working with the South San Francisco Unified School District. In 1974, while looking for a more efficient way to provide reading tutoring for special education students, he discovered a Stanford University project on computer-assisted instruction, which was adapted to support a phonics tutoring project called Success Controlled Optimal Reading Experience (SCORE).



Ruthmary Cradler, vice president of ESS, works as a consultant, specializing in evaluation of Ed Tech program implementation and development and is a coordinator on the CARET project. She has helped state and local agencies develop technology and No Child Left Behind plans. A member of the PT³ national advisory committee for program evaluation, she holds an MA from San Francisco State University.



Robert Clarke is a retired professor of psychology, measurement, and statistics from San Jose State University. He holds a PhD from Stanford University. His experience in education began in the late 1950s and early 1960s as a high school math teacher and a research consultant for school districts.