

# **Evaluating What Works in Education: Causation or Context.**

AEA 2009 Paper Presentation  
by  
Randall Davies  
Brigham Young University

## **Citation:**

Davies, R. (2009, Nov). *Evaluating What Works in Education: Causation or Context*. Paper presented at the American Evaluation Association's annual conference, Orlando, FL.

## **Abstract**

Is it possible to NOT incorporate context into our evaluation inquiries? When attempting to determine best practices and what works in educational program evaluations, are the questions causal or contextual? This paper argues that in educational research and evaluation, often what we assume to be causal questions are actually by nature contextual. Understanding the context of our inquiries is the foundation for rigorous and valid observations. Understanding the true nature of the learner in the teaching and learning process is essential to any evaluation of educational programs.

## **Introduction**

In program evaluation situations we commonly hear the terms proven practices, research based programs, or “what works.” For some, identifying educational practices that cause learning to occur is the ultimate objective of educational research. In fact, an expressed purpose of the current Elementary and Secondary Education Act, commonly referred to as No Child Left Behind (NCLB), is to raise achievement for all students and to close the achievement gap through the expectation that schools use proven educational methods (U.S. Department of Education, 2002). The causal assumption being that if researched-based proven practices are used by teachers then students will learn. As a result, in 2002 the What Works Clearinghouse was established as a repository of scientific evidence for what works in education (What Works Clearinghouse, 2002).

Unfortunately, it is quite possible that research-based proven practices are a myth. They do not exist. The quest for educational practices that are proven to work in schools while admirable may be misguided; or rather, we may be looking in the right place but at the wrong thing. When looking for what causes learning to occur we must consider the fact that while educational programs can effectively facilitate learning under certain conditions, they do not

cause learning to occur. And while a well-substantiated body of research links teacher knowledge and experience with student learning (Darling-Hammond, 2000; Goldhaber & Brewer, 1996; Mendro, 1998; Stedman, 1997; Wenglinsky, 2002), being labeled highly qualified and using research-based methods does not mean the teachers will be effective for all students in all situations. Despite the best efforts of excellent teachers, many students do not attain proficient levels of achievement, and even in the best classrooms, not all students achieve excellence; some even fail (Keller & Susuki, 1988). This is because learning is not the result of so call proven programs but rather the result of individual students' efforts to construct their own knowledge, understanding, and ability (Woolfolk, 2008).

When an evaluator wrongly assumes that the evaluation of an instructional intervention is a question of causation, the selection of appropriate tools may be a problem. When interpreting evaluation data, again, the issue of context is paramount. Without a proper understanding that the program is part of the context, evaluation results may lack interpretive validity. This paper outlines an argument for how selection of evaluation methods and the interpretation of evaluation results must be informed by a proper understanding of context and the true nature of the learner as the cause of learning when evaluating educational programs.

## **Background**

In the not so distant past, member of the American Evaluation Association (AEA) experienced a civil war. It started with a debate over quantitative and qualitative points of view. This escalated into a theoretical argument regarding methods; which culminating in a funding priority proposed by the U.S. Department of Education (2003) that proclaimed evaluation methods using an experimental design were best for determining an educational program's effectiveness. The AEA's response to this priority focused on the issue of cause and effect, as well as the issue of scientifically rigorous methods (American Evaluation Association, 2003). A reconciliation of sorts was obtained when most AEA members seemed to acknowledge the value of alternative methods, with the concept of a mixed methods approaches to evaluation forming a common ground. Yet while the war has subsided, the issues remain largely unresolved. I believe this is largely due to the fact that many researcher and evaluators inaccurately identify educational programs as the cause of learning rather than a part of the context. This misconception tends to promote a wrong view of the relationship between instructional programs, the teacher, and the learner in the teaching and learning process. Misunderstanding the teaching and learning process affects how we evaluate schools as well as the policies and legislation we might propose.

## **Conditions for Establishing Cause and Effect**

Research cannot and does not attempt to prove whether an educational intervention will work in all situations and contexts (Johnson & Christensen, 2004; Shadish, Cook, & Campbell,

2002). Researchers simply attempt to provide evidence of the potential effectiveness of a program within the confines of a specific context. And yet we still claim that research can identify proven practices. This may be a simple issue of defining terminology but it is an important point to understand. While many attempt to pose the proven practices issue as a cause and effect question, it is not. If we were to look at the three conditions for establishing a cause and effect relationship (i.e., existence of a correlational relationship, temporal antecedence, and no extraneous confounding factors), no educational practice meets these conditions. While there may exist a correlation between a beneficial practices and learning, learning can and does occur without the treatment or intervention. Many students learn the intended content prior to being instructed. More importantly, learning occurs at different rates and in different degrees for different students given the same intervention. This suggests, as educational psychology theory states, the cause of learning is not the instructional intervention but rather some other factor (Woolfolk, 2008). When evaluating educational programs in schools that extraneous causal factor is the individual student.

**The monotonicity assumption.** In addition to the failure to meet the conditions for establishing educational practices as causal agents, there is another indicator that educational programs alone do not cause learning to occur. The monotonicity assumption contends that in order to produce an unbiased estimate of a causal relationship there must not be any defiers in the studies comparison groups (see Angrist, Imbens, & Rubins, 1996). This insight acknowledges the fact that, in real life classroom situations, there are and will always be those who do not comply with treatment. In other words, the issue of agentive intentionality must be considered when evaluating the effectiveness, merit or worth, of any educational program. The fact that not all students will cooperate or comply with treatment is an unavoidable condition of the educational context. When conducting a program evaluation, noncompliance, disinterest, or lack of effort by some students should not be interpreted as evidence of a failed intervention. Most educational interventions may, and in most cases do, benefit compliant participants to some degree. Reading a textbook facilitates learning if the student is intent on gaining knowledge, reads well, and has access to the materials. Yet it is not the intervention that causes the learning to occur, rather the intervention facilitates the learning that is caused by the compliance students' attention to the facilitating activity. Learning occurs within the learner and is caused by the learners' attending to the intended learning activity.

### **Impact with Attribution**

A recent adaptation of terminology being used to describe the cause and effect relationship is the term impact with attribution. This is a softer way to say a specific outcome (i.e., the impact) can be attributed in some way to an educational program, practice, or instructional intervention (i.e., the likely cause). It basically means the same thing but leaves room for the fact that the effect may not occur in every instance due to uncontrollable confounding factors. It also acknowledges an understanding that effective programs, although

not the cause, may contribute to a student learning what was expected of them, in that compliant learners may benefit to some degree from participating in the activity.

### **Conditions for Learning**

Constructivist theories of learning in educational psychology suggest that it is the individual student that is the cause of learning (Woolfolk, 2008); but contextual factors have a huge impact in the way and the degree to which students learn (Bronfenbrenner, 2005). Learning happens everywhere. We do not need to go to school to learn but schools attempt to focus learning on specific learning objectives. Educational programs and practices are intended to help student learn what is expected of them in that designated curriculum.

Three conditions must exist for intended learning to occur in schools: (1) students must have sufficient cognitive ability and they must be developmentally ready to learn, (2) students must be provided with the opportunity to learn (i.e., be exposed to ideas, have access to resources and training), and (3) students must be willing to learn what is expected of them. Of these three conditions educational research tends to focus on evaluating the opportunities provided. However, providing students with a quality learning opportunity is no guarantee that students will learn. All three conditions must exist.

### **Understanding the True Nature of the Learner**

In describing his experience on the What Works Clearinghouse technical advisory group at the 2005 annual AEA conference, William Shadish lamented what he called the problem of “dueling experts.” He commented that he was sure we could figure out how to make it work if only we could remove the human element from the mix. During the question and answer portion of the presentation Michael Patton rose and suggested that for some evaluators the human element was the most interesting part. He might have added the only part worth understanding.

Clearly the complexity of the human element makes evaluating any educational program challenging. Having an opportunity to learn is important, as is a student’s ability, but only when we correctly understand the true nature of the learner within a teaching and learning context can an evaluation be conducted properly. A valid complaint leveled against many educational researchers, instructional designers, and policy makers is that they cling to a wrong minded, fundamentally flawed view of the student in the teaching and learning process (Gordon & Zemke, 2000). A revised view of the student would replace the discrete and largely passive models of average student functioning with models that account for differences in student’s goals and intentions, knowledge about themselves and the environment, and an individual’s ability to develop and change strategies of action (Renninger, Hidi, & Krapp, 1992). A proper view of the student is one of intentionality and agency. It takes into account a learner’s intent, their commitment and their desire to learning what is expected of them when presented with a specific learning opportunity; it accounts for their motivation to learn and the effort they are willing to

put forth (Davies, 2009). When we insist on identifying educational programs as causal agents, rather than the individual students, we can and often do misinterpret the results of our inquiries.

### **Implications for Evaluation of Educational Programs**

If educational programs do not cause learning to occur then the concept of proven practice is a myth; the goal of educational research to identify proven practices is misguided; and the accountability mandates we place on educational professionals to use only research-based proven practices are unfounded. A proper understanding of educational programs as part of the educational context and not the cause of learning allows us to think differently about the purpose of educational research and evaluation. It also allows us to think differently about the methods we use to evaluate the effectiveness of an educational program.

### **Educational Programs and the Goal of Educational Research**

While it may be a small thing, the use of the term “proven practice” should not be used when referring to what works in education. In a real sense, educational programs are simply a part of the context in which learning can occur. Educational programs may facilitate learning but they do not cause the learning. An educational program is properly seen as part of the environment in which learners accomplish their learning goals. Recognizing individual students as causal agents, an important goal of educational research in this regard should be to identify promising practices that might enhance a specific learning environment. The degree to which an educational program is deemed effective is then understood as the degree to which capable compliant learners benefit from the program, intervention, or practice. We would expect pedagogically sound practices be used regardless of whether student test score change.

### **Methods for Evaluating Educational Programs**

When evaluating educational programs it is common to look at whether the educational practices being used are research-based. When used as an adjective describing educational practice, the research-based label is intended to denote the notion that quality research was conducted and the results support the conclusion being made. Unfortunately, being research based could mean almost anything. Research comes in many forms and serves a variety of purposes. The decision to include a study as evidence of what works in the archives of the What Works Clearinghouse is based primarily on the research methods used (What Works Clearinghouse, 2002). The gold standard for research designed to determine program effectiveness, and thus what practices teachers should use in schools, is an experimental design (U.S. Department of Education, 2003). The reason for this is the belief that an experimental design is the strongest research method for identifying generalizable causal relationship (Gay, 1996; Johnson & Christensen, 2004; Shadish, Cook, & Campbell, 2002; Slavin, 2007). Unfortunately, educational programs are not causal agents. The fact that a program is not the cause of learning may be the reason why program evaluation using random controlled trials to

date have failed to revolutionize the effectiveness of the education process to any great extent (Branson & Hirumi, 1994; Design-Based Research Collective, 2003).

Experimental methods can be useful in determining the degree to which a specific intervention may have contributed to average group test scores compared to a control group as measured by a specific testing instrument at a specific moment in time within a specific context; but in the complex and dynamic world in which students and educators exist, this type of information may be of little practical value to educational practitioners and society in general (Shadish, Cook, & Campbell, 2002). Program effectiveness determined in this way falls into the category of what Stufflebeam and Shinkfield (2007) calls quasi-evaluation because it focuses narrowly on only one context dependent indicator of effectiveness. In practice, using a variety of qualitative and quantitative methods that look at a variety of criteria in many settings by multiple evaluators may be the only and most reasonable method for determining promising practices or what truly works in education (Davies, Williams, & Yanchar, 2008).

### **Accountability Expectations and Evaluating Educational Programs**

One of the most problematic aspects of NCLB accountability expectations is the limited way in which teacher and schools are evaluated. For teachers, using appropriate instructional practices is important but not as important as a teacher's ability to implement a program well. It is simply not enough to say research-based programs are being used. Neither is it appropriate to judge a teacher's efforts to provide pedagogically sound instruction within a beneficial learning environment solely on test scores (Davies, 2008). The true nature of the learner and their place in the teaching and learning process is an important factor for evaluators to consider. Teacher should care about their students and do what they can, within reason, to help them learn. It is the students' responsibility to take advantage of the opportunities they are given. Students should be held accountable for their own achievement. Assessment results help teachers and students identify areas for improvement but test scores should not be used as the sole criteria for teacher evaluations or determining whether an educational program is pedagogically sound.

### **Conclusions**

Researchers and evaluators at times inaccurately identify educational programs as the cause of learning. This misconception tends to promote a wrong view of the relationship between instructional programs, teachers, and learners in the teaching and learning process. Misunderstanding the teaching and learning process often adversely affects how we evaluate school programs. What we call research-based proven practices do not exist if by proven practice we mean the program is guaranteed to work. In addition, using research methods designed to identify causal relationships may provide insufficient evidence of program effectiveness. A rigorous evaluation of a program would require much more than an assessment of student test scores (Davies, 2008; Fitzpatrick, Sanders, & Worthen, 2004).

A proper understanding of an educational program is to see it as a part of the learning context. Educational programs and practices form a part of the learning environments teachers try to foster in the hope that these promising practices will facilitate learning. When conducting an evaluation, understanding the context of our inquiries is the foundation for rigorous and valid observations. We need to recognize that no educational program or practice will be effective with non-compliant students. Most educational interventions have merit and worth for students who's intention is to learn what is expected of them. Certainly not all teachers implement pedagogically sound practices well; still, one problem with current accountability mandates is the erroneous assumption that good teaching always results in increased academic achievement and learning. Learning is a function of a student's actions and efforts. It can be facilitated by good teaching, but instruction is not the casual factor. Understanding the true nature of the learner in the teaching and learning process is essential to any evaluation of educational programs.

## References

- American Evaluation Association (2003). Response To U.S. Department of Education Notice of proposed priority, Federal Register RIN 1890-ZA00. Retrieved from <http://www.eval.org/doestatement.htm>
- Angrist, J. D., Imbens, G.W., & Rubins, D. B. (1996). Identification of causal effects using instrumental variables (with discussion). *Journal of the American Statistical Association*, 91, 444-472.
- Bronfenbrenner, U. (2005). *Making human beings human: Bio-ecological perspectives on human development*. Thousand Oaks, CA: Sage.
- Branson, R. & Hirumi, A. (1994). Designing the future: The Florida Schoolyear 2000 Initiative. In G. Karkly & W. Lynch (Eds.) *Educational technology: Leadership perspectives* (pp. 91-112). Englewood Cliffs, NJ: Educational Technology Publications.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, 8(1). Retrieved from <http://epaa.asu.edu/epaa/v8n1>.
- Davies, R. (2008) AYP accountability policy and assessment theory conflicts. *Mid-Western Educational Researcher*, 21(4), pp. 2-8.
- Davies, R. (2009). *Exploring the Meaning and Function of Learner Intent*. Saarbrücken, Germany: VDM Publishing House Ltd.
- Davies, R., Williams, D., & Yanchar, S. (2008). The Use of Randomization in Educational Research and Evaluation: A critical analysis of underlying assumptions. *Evaluation & Research in Education*, 21(4), pp. 303 – 317.
- Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5-8.
- Fitzpatrick, J., Sanders, J., & Worthen, B. (2004). *Program Evaluation: Alternative Approaches and Practical Guidelines* (3rd edition). Boston, MA: Pearson Education, Inc.
- Gay, L. R. (1996). *Educational research: Competencies for analysis and application* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Goldhaber, D. D., & Brewer, D. J. (1996). Evaluating the effect of teacher degree level on educational performance. *ERIC Digest*. (ERIC Document Reproduction Service No. ED406400).
- Gordon, J., & Zemke, R. (2000). The Attack on ISD. *Training*, 27, 43-53.



- Johnson, B. & Christensen, L. (2004) *Educational research: Qualitative, quantitative, and mixed methods approaches*. (2nd ed.). Boston: Pearson Education.
- Keller, J. M., & Susuki, K. (1988). Use of ARCS motivational model in courseware design. In D. H. Jonassen (Ed.), *Instructional designs for microcomputer courseware*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mendro, R. L. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personal Evaluation in Education*, 12, 247–256.
- Reigeluth, C., & Beatty, B. (2003). Why children are left behind and what we can do about it. *Educational Technology*, 43(5), 24-32.
- Renninger, K. A., Hidi, S., & Krapp, A. D. (1992). *The role of interest in learning and development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Slavin, R. (2007). *Educational Research in an age of accountability*. Boston, MA: Pearson Education Inc.
- Stedman, L. C. (1997). International achievement differences: An assessment of a new perspective. *Educational Researcher*, 26, 4–15.
- Stufflebeam, D.L. & Shinkfield, A.J. (2007). *Evaluation Theory, Models, & Applications*. San Francisco, CA: Jossey-Bass.
- U.S. Department of Education, Office of Elementary and Secondary Education (2002). *No child left behind: A desktop reference*. Retrieved from <https://www.ed.gov/admins/lead/account/nclbreferenc/page.html>
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (2003). *Identifying and implementing educational practices supported by rigorous evidence: a user friendly guide*. Retrieved from <http://www.ed.gov/rschstat/research/pubs/rigorousetid/rigorousetid.pdf>
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(12). Retrieved June 20, 2008 from <http://epaa.asu.edu/epaa/v10n12/>.
- What Works Clearinghouse (2002). *About Us statement*. Retrieved from <http://ies.ed.gov/ncee/wwc/aboutus/>
- Woolfolk, A. (2008). *Educational psychology*, Tenth Edition: Active Learning Edition. Pearson Education, Inc.