

# The Role of External Evaluators in Educational Projects and Programs

B.M. Moskal\*

*Mathematical and Computer Sciences Department, Colorado School of Mines, Golden, Colorado, USA*

**Abstract:** This article examines the benefits of including a trained external evaluator on educational research projects and programs. The argument is presented that, when projects or programs impact educational policy or decision making, assessment and evaluation should be guided by a systematic plan implemented by a trained evaluator.

**Keywords:** Assessment, evaluation, education methodology.

## INTRODUCTION

Currently, I am a Professor in the Mathematical and Computer Sciences Department and the Director of the Center for Assessment of the Science, Technology, Engineering and Mathematics (STEM) disciplines at the Colorado School of Mines (CSM), Golden, Colorado, USA. For more than ten years, I have acted as an evaluator for educational research projects and programs in STEM in the United States. My evaluation efforts have spanned elementary school, middle school, high school and college levels, in fact, the entire educational spectrum. I separate educational research projects from educational programs because of the different purposes that they serve. Educational research projects seek to answer a question for which there is a void in the current literature, whereas educational programs often use research to inform the development and implementation of a program but their purpose is not necessarily to expand current research. Both, however, require quality assessment and evaluation to inform their goals and improve their approaches.

The purpose of this article is to clarify the role of external evaluators on educational research projects and programs. There are two events that inspired the writing of this article. First, I was recently asked to review an article concerning assessment written by scientists who were implementing an educational program. I was surprised to read that, based on the opinions of the authors, the purpose of hiring an external evaluator is solely to satisfy the requirements of funding agencies. Reading further, I learned that these authors believe that the development and implementation of an appropriate assessment and evaluation plan in such programs requires neither technical expertise nor an evaluation background. Their message was clear: if you are not required to hire an external evaluator, save your money. They explain further that faculty and K-12 teachers have the background necessary to implement a quality assessment and evaluation plan for educational research projects and programs as a result of the experiences that they have evaluating students' knowledge in the classroom. Although contrary to my

recommendation the article is scheduled for publication, I do not reference it here in consideration of its authors and my negative evaluation of the article's contents.

The second event that motivated the current article occurred at a conference focusing on assessment and evaluation, which I was unable to attend. One of my graduate students was to give a presentation on his research which involved assessment and evaluation. Upon the student's return from the conference, I was shocked to learn that his talk was not well received and that he had used only a fraction of his allotted time. He went on to explain that a well meaning member of the faculty from another university who had no assessment or evaluation experience but who did have a Ph.D. in a science field had reviewed his talk on the evening before the presentation. The student agreed to this review based on a discussion with this faculty member during the conference. The faculty member had attended a few sessions on assessment earlier in the conference and decided that the student's assessment was not designed correctly. In order to save the student embarrassment, he recommended that the student remove the sections that he felt were problematic, leaving the student with little to present. Through a phone call, I confirmed this sequence of events with the faculty member.

Wishing to see the research that inspired the faculty member's behavior, I systematically reviewed the websites and the papers that he claimed challenged the assessment design. The research that he identified had only face validity with respect to the student's project. This is not to say the research was not well done but rather it was not related to the student's investigation. None of the research addressed the same content and constructs as the student's talk nor was it completed with comparable student populations. Some of the methods were interesting but they did not challenge the approach that was used. The faculty member's actions did not prevent embarrassment but rather he caused it.

These examples illustrate a major concern in assessment and evaluation. People who have been assessed as students and currently assess their own students may believe that these experiences qualify them to judge, develop, and implement broad scale assessment programs. Many of these individuals are unknowingly ill prepared for the task at hand.

## FALLACY: ASSESSMENT AND EVALUATION REQUIRE NO SPECIAL TRAINING

Taken at face value, the claim that assessment and evaluation require no special training is true. Anyone can assess;

\*Address correspondence to this author at the Mathematical and Computer Sciences Department, Colorado School of Mines, Golden, Colorado, USA; E-mail: bmoskal@mines.edu

anyone can evaluate. Individuals who have had no training in assessment and evaluation can collect, analyze and interpret data. The fallacy is the belief that the quality of the information acquired through these activities is likely to be high.

I illustrate my point with an analogy. Much in the same way that anyone can assess, anyone can do science. Young children, for example, test gravity by throwing items from their high chairs. They may even do this repeatedly. Participating in a scientific experiment or even multiple scientific experiments does not qualify a person as an expert in the field.

This common sense argument in science is often ignored in assessment and evaluation. People who have participated in assessment may believe that they are qualified to develop, implement, interpret and evaluate assessments as well as make broad scale recommendations based on assessment outcomes. This problem was illustrated through the two examples provided in the introduction to this article. The authors of the article that I reviewed did not see the value of an evaluator because they believed that their classroom assessment efforts qualified them as evaluators. After attending a few sessions on assessment, the well intending faculty member believed that he too was an expert on assessment and provided poor advice to the graduate student.

The above argument is not to imply that it is impossible to acquire quality information from an untrained evaluator. Certainly, there are exceptions. In general, however, it is poor policy to base educational decisions on the recommendations of untrained evaluators.

## PHASES FOR ASSESSMENT AND EVALUATION

There are several well recognized phases for the assessment and evaluation process [1-5]. First, a plan needs to be established for the collection and analysis of data. The quality of this plan directly impacts the quality of the conclusions that can eventually be drawn. In order to develop such a plan, the investigators should have a clear understanding of what it is that they are seeking to measure [6]. Therefore, the planning phase begins with the articulation of project or program goals, objectives and outcomes. Outcomes should be stated in a manner such that they can be measured. Even when the goals of the program concern psychological constructs, the outcomes should be designed to reflect external manifestations of those constructs. The process of articulating project or program goals in a manner in which they can be measured is well understood by most evaluators.

The second phase is to develop a research methodology that may be used to collect the appropriate information. Research methodology refers to the strategy that is used to obtain interpretable data. Graduate courses are offered in schools of education that address both quantitative and qualitative methodological approaches. By understanding the benefits and drawbacks of various methodological approaches, an evaluator can assist in the selection and implementation of methods that are likely to yield interpretable results [7].

A major component of the research methodology is the development or selection of assessment instruments. A common mistake on the part of many educational projects is the use of internally developed instruments which have not been validated. The use of such instruments results in evidence that can only weakly support conclusions. Validated instruments that are

designed to measure an array of content and constructs are available through a search of the Educational Testing Service Testlink [8] or the ERIC database [9]. Evaluators within a given field are often aware of other services available to that discipline. When a validated instrument cannot be found for a given set of outcomes, an evaluator can assist the investigator in developing an internal instrument and guide the validation of that instrument, an intense and time consuming process.

The third phase of the assessment process is the collection of data. Before this phase can be completed in the USA, the project or program should seek and receive the approval of an Internal Review Board (IRB). An IRB is a group of faculty members who receive training to determine whether a research investigation is likely to cause harm to participating subjects. The United States Department of Health and Human Services provides a variety of resources concerning the regulations that surround the collection of data from human subjects in the USA [10]. Many investigators incorrectly assume that if their work "causes no harm" based on their own judgment then the work is exempt from IRB oversight. The decision that a study is exempt should not be made by an investigator but rather by a third, unbiased party, such as an IRB. An IRB is qualified to judge exemptions. An evaluator who has previously completed this process can assist investigators in navigating the IRB process.

Another component of the assessment process is how the data are collected. If data are to be collected across multiple classrooms, then attention should be given to the process of ensuring consistency in the administration of the instruments and the collection of the data across classrooms. Deviations that occur in the administration can result in the introduction of external, uncontrolled factors and could render the results difficult or impossible to interpret. Evaluators can assist in the development of a process that ensures the consistent administration of instruments and collection of data.

The fourth phase of the assessment process is the analysis of data. This phase is guided by the assessment plan established in the first phase. If the data are to be analyzed statistically, the evaluator can prepare the data for analysis purposes, evaluate the extent to which the data conforms to the statistical assumptions, and complete the analyses. If the assumptions of a statistical test are not met, an alternative analysis may be proposed and implemented. An evaluator who has studied analysis techniques and has applied them in various settings is equipped to complete this type of analysis. Statisticians are also often consulted during this phase. A valid concern of many statisticians with regard to educational research is the lack of random assignment; schools in the USA rarely have the option of randomly placing students into treatment and control groups [11]. Evaluators understand this complexity of educational research and are often well versed in quasi-experimental design methods, which recognize that random assignment is often not feasible [12]. In other words, evaluators can propose alternative statistical techniques with the purpose of controlling for external factors.

Based on the specified outcomes of an educational project or program, the most compelling evidence to address a given set of outcomes may be qualitative data [13]. Many investigators confuse qualitative data with anecdotal evidence and, therefore, incorrectly assume that they are well versed in qualitative techniques. Anecdotal evidence consists of observations that have been made in an unsystematic manner; qualitative analysis

is guided by a systematic research plan. There are specific methods that guide the collection and analysis of qualitative data which help to ensure the trustworthiness of the interpretations of that data. Evaluators can assist in the development of qualitative research plans, and in the collection, analysis and interpretation of the data.

Some evaluators specialize in quantitative analysis, others specialize in qualitative analysis, and still others have knowledge of both. Even when evaluators do not have the appropriate knowledge of a given form of analysis, they typically are aware of consultants who have the complementary skill base.

The final phase of the assessment process is the interpretation and use of results. This is the "evaluation" portion of the assessment and evaluation process. The extent to which the results provide meaningful information is often dependent on the appropriateness of the original methodology. Occasionally, well designed assessment programs fail to provide compelling results. However, poorly designed assessment programs rarely provide compelling evidence. Whenever possible, it is best to plan for the collection and analysis of data in order to avoid the collection of data that has no meaning. A quality evaluator can guide this process.

### ROLE OF AN EVALUATOR

An evaluator can enter the assessment process during any of the above described phases [6]. It is typically best, however, to include an evaluator from the start of the project or program. This is especially important when the investigators do not have an assessment and evaluation background. By including an evaluator at the beginning, the evaluator can assist in the development of appropriate goals, objectives and outcomes. Insight can be provided by the evaluator to the investigators as to what can reasonably be measured. The evaluator can further lead the effort to develop and implement the assessment plan and complete the final analysis.

When an evaluator is recruited to a project or program that is already underway, problems are likely to emerge. Without a systematic plan from the beginning, the collected data, regardless of whether they are qualitative or quantitative in form, may be difficult or impossible to interpret. Furthermore, the goals, objectives and outcomes of the project may be poorly defined or not defined. An evaluator cannot correct a poorly designed plan after the data have been collected. Many investigators are also unaware of the impact that external factors can have on their results. Without consideration and control for these factors, interpreting the final results can be a challenge. Serious flaws in the research design or the data collection process threatens the validity of the interpretation of the results. The inclusion of an evaluator in the early phases of a project or

program can minimize the impact of these factors on the outcomes of the project or program.

### CONCLUSIONS

I have heard it argued, "You don't need a Ph.D. to do assessment." I would like to provide my support for this statement with a caveat; you also do not need a Ph.D. to do science. However, as a society we rarely base our scientific decision making on the discoveries of individuals who have not been trained in science. As a society or as a discipline, do we really want our educational decisions to be based on the results of untrained evaluators? Assessment and evaluation, when done well, *do* require technical skills. There is a process for collecting, analyzing and interpreting data and there are many factors that can impact the outcomes. Without evaluators who have an appropriate background, the quality of the interpretations that are made of such data is likely to be of questionable value.

### REFERENCES

- [1] Committee on the undergraduate program in mathematics. Assessment of student learning for improving undergraduate major in mathematics. [Accessed 29 May 2008]. Available from: <http://www.maa.org/saum/maanotes49/279.html>.
- [2] McGourty J, Sabastian C, Swart W. Developing a comprehensive assessment program for engineering education. *J Eng Educ* 1998; 87 (4): 355-61.
- [3] Moskal B. Using assessment results for improving student learning. In: Spurlin J, Rajala S, Lavelle J, Eds. *Designing better engineering education through assessment: a practical resource for faculty and department chairs on using assessment and ABET criteria to improve student learning*. Virginia, USA: Stylus Publishing 2008: pp. 117-48.
- [4] Rogers G. Death by assessment. [Accessed 29 May 2008]. Available from: <http://www.abet.org/Linked%20Documents-UPDATE/Assessment/Assessment%20Tips2.pdf>.
- [5] Shadish W. Some evaluation questions. [Accessed 29 May 2008]. Available from: <http://PAREonline.net/getvn.asp?v=6&n=3>
- [6] Gajda R, Jewiss J. Thinking about how to evaluate your program? These strategies will get you started. [Accessed 29 May 2008]. Available from: <http://PAREonline.net/getvn.asp?v=9&n=8>.
- [7] Moskal B, Leydens J, Pavelich, M. Validity, reliability and the assessment of engineering education. *J Eng Educ* 2002; 91 (3): 351-4.
- [8] Educational Testing Service. "Test link." [Accessed 29 May 2008]. Available from: <http://www.ets.org/testcoll>
- [9] ERIC. SearchERIC. [Accessed 29 May 2008]. Available from: <http://searcheric.org>.
- [10] United States department of health and human services. Office of human research protection. [Accessed 29 May 2008]. Available from: <http://www.hhs.gov/ohrp/>.
- [11] Olds B, Moskal B, Miller R. Assessment in engineering education: evolution, approaches and future collaborations. *J Eng Educ* 2005; 94 (1): 13-25.
- [12] Gribbons B, Herman J. True and quasi-experimental designs. [Accessed 29 May 2008]. Available from: <http://PAREonline.net/getvn.asp?v=5&n=14>.
- [13] Leydens J, Moskal B, Pavelich M. Qualitative methods used in the assessment of engineering education. *J Eng Educ* 2004; 93(1): 65-72.