# Evaluation of Educational Development Projects*

*This paper published at the ASEE/IEEE FIE2006 Conference*

Russell Pimmel[1], Barbara Anderegg[2], Susan Burkett[3], Bevlee Watford[4], and Connie Della-Piana[5]

*Abstract* – **This paper is a narrative form of an interactive session that aims to develop a better understanding of the evaluation process and the tools and techniques used in such a process. This enhanced understanding will enable you to work with an expert evaluator in preparing and implementing an evaluation plan for an education development project. The paper focuses on the role of goals and outcomes in the evaluation process, on the nature of various types of outcomes, on evaluation tools for monitoring these types of outcomes, and on the issues involved in the interpretation of evaluation data. It suggests several issues to consider when writing an evaluation plan and when working with an evaluator.**

*Index Terms* – assessment, goals and outcomes, evaluating outcomes, evaluation tools, interpreting evaluation data, evaluation plans

## INTRODUCTION

Engineering educators are becoming increasingly involved in both funded and unfunded projects to improve student learning, success, and diversity. Monitoring the progress of these efforts, identifying their positive and negative impacts, and ultimately determining their success or failure requires systematic evaluation efforts. Since most engineering faculty members have little understanding of this activity they hesitate to undertake this step. Even if they plunge ahead they do an inadequate job, especially if they wait until the last few months of the project. Even if they decide to seek help from experts in the evaluation field, the lack of understanding on both sides complicates the interaction and often leads to an uninformative evaluation. In order to try to improve this situation, several program directors in NSF's Division of Undergraduate Education developed an interactive session or workshop on project evaluation to provide engineering faculty with a framework for thinking about this topic and to provide the essential background to allow them to more effectively discuss evaluation with experts in the field.

The original interactive workshop utilized a "think, share, report, and learn" format where the structure of the activities encouraged the participants to recall their current knowledge, identify and correct misconceptions, and connect new knowledge to their existing knowledge. This involved a number of exercises where participants first formulated their own response, shared these in a small group and then in a large group, and then compared their response to that of an "expert." In this case, the "expert" was the consensus of a group of NSF program directors. In order to make this material more widely available, we converted it to a narrative format resulting in this paper. In converting this to a narrative format, we have tried to retain an engaging format by including the same exercises and then asking the reader to formulate his or her response before considering that of the "experts."

The goal of this tutorial paper is to help you develop a better understanding of the evaluation process and the tools and techniques used in such a process. This enhanced understanding will enable you to work more effectively with an expert evaluator in preparing an evaluation plan for an education development project and to actually implement an evaluation of an existing project. This enhanced capability will allow you both to prepare more competitive proposals for NSF's education programs and to determine the effectiveness of all educational development efforts regardless of the funding source. In pursuit of this goal, the material intends to increase the participants' awareness of the role of goals and outcomes in the evaluation process, of the nature of various types of outcomes, of evaluation tools for monitoring these types of outcomes, and of the issues involved in the interpretation of evaluation data. In subsequent sections, the paper will address the following six issues:

- Using goals and outcomes in project evaluation,
- Evaluating cognitive outcomes,
- Evaluating affective outcomes,
- Interpreting evaluation data,
- Writing an evaluation plan for a proposed project,
- Working with an evaluator.

Since the term "evaluation" and, in the same way "assessment," are used in many contexts with slightly different meaning in each, we need to define ours before proceeding. In considering evaluation in engineering education, there are at least three common contexts: evaluating individual performance or grading, evaluating program effectiveness as

* The ideas presented in this paper represent the views of the authors and not an official policy of the NSF.

[1] Russell Pimmel, Program Officer, Division of Undergraduate Education, National Science Foundation, rpimmel@nsf.gov
[2] Barbara Anderegg, Program Officer, Division of Undergraduate Education, National Science Foundation, bandereg@nsf.gov
[3] Susan Burkett, Program Officer, Division of Undergraduate Education, National Science Foundation, sburkett@nsf.gov
[4] Bevlee Watford, Program Officer, Division of Undergraduate Education, National Science Foundation, bwatford@nsf.gov
[5] Connie Della-Piana, Program Evaluation Manager, Office of Integrative Activities, National Science Foundation, cdellapi@nsf.gov

in the ABET assessment process, and evaluating a project's progress and success. This material addresses the third, project evaluation. Even though the three evaluation contexts utilize many of the same tools and processes, the different goals result in important differences. Project evaluation, the focus of this paper, may involve either individual evaluations as in grading, or group evaluation as in the ABET process, or both – but all within the context of the project.

## GOALS, OBJECTIVES, OUTCOMES, AND EVALUATION QUESTIONS

Evaluation starts with carefully defined project goals that lead to measurable outcomes, and these in turn lead to evaluation questions. Many of a project's goals and outcomes will relate to project management, for example those associated with initiating or completing a project activity and those associated with finishing a "product." However, in an education development project, some of the goals and outcomes should relate to changes in student behavior, for example by modifying student learning, attitudes, or perceptions.

In the beginning, when the project team begins to define the project, they should identify one or more overarching statements of the project's intention or ambition, which in this discussion we will call these project *goals*. Recall that the intent of this paper is to enable you to work with an evaluator, and so the project team should include this expertise as it works to define the project's goals.

Early in the development, the project team should transform or convert each goal into one or more expected measurable results and we define these statements of the expected measurable results as *outcomes.* If the goal is very general, describing a broad intention, it is sometimes convenient to proceed through one or more intermediate steps where these intermediate statements of intentions become more specific and more measurable than goals but less specific and less measurable than outcomes. In this discussion we will refer to these intermediate statements as *objectives*. Converting very general goals into outcomes may require more than one intermediate step, leading to, for example, "high-level objectives " and "low-level objectives".

As the project team begins to think about evaluation, the outcomes lead into evaluation or, in some sense, research questions. Suppose that a project involves some change in classroom activities, which we will refer to as an *intervention*. As a result of the intervention, the team expects that student behavior, reflected, for example, in learning outcomes, attitudes, or perceptions, will be different than those seen before the intervention. These changes are the expected outcome. Two types of evaluation questions arise – one dealing with whether or not the change occurred and one dealing with whether or not the intervention actually caused the change. The evaluation process should address these questions by collecting, analyzing, and interpreting data on the observed behavior and on the factors that could cause this kind of change.

We now present the material from the first few exercise in the workshop, which deal with defining project goals, transforming goals to objectives and outcomes to evaluation questions. You may treat these as extended examples or as interactive exercises that you complete before proceeding to the "expert's" response. We hope that you will approach it as exercise because we believe that actively engaging in the process will increase your learning, and so we urge you to take a few minutes and write your responses to each exercise.

***Exercise #1 – Writing Goals:*** *The partial abstract shown in Figure 1 is a modified version of an abstract for a successful NSF proposal for a $150,000 award. We have deleted the statement of the project's goals, along with the details of the evaluation and dissemination plans. Before proceeding to the program directors' responses, you should write a few plausible goals for this project with at least one of them focused on a change in learning and one on a change in some other aspect of student behavior.*

**Modified Abstract for NSF Project**

**The goal of the project is … The project is developing computer-based instructional modules for statics and mechanics of materials. The project uses 3D rendering and animation software, in which the user manipulates virtual 3D objects in much the same manner as they would physical objects. Tools being developed enable instructors to realistically include external forces and internal reactions on 3D objects as topics are being explained during lectures. Exercises are being developed for students to be able to communicate with peers and instructors through real-time voice and text interactions. The material is being beta tested at multiple institutions including community colleges. The project is being evaluated by … The project is being disseminated through …**

FIGURE 1
MODIFIED ABSTRACT

***Response #1 – Writing Goals****: In this response, and in all those that follow, we list several suggestions made by NSF program directors; you should consider these as examples and not as a complete list of all possible responses and certainly not as an official NSF response. In this case, the NSF program directors identified a number of different types of goals and we have divided these into four categories, dealing with changes in cognitive behavior, affective behavior, success rate, and diversity.*

*Goals in the first category, those related to changes in cognitive behavior, fit into two groups. The first relates to concepts or applications explicitly taught in the course, and examples include goals to increase the students' ability to:*

- *Solve textbook problems,*
- *Draw free-body diagrams for textbook problems,*

- *Describe verbally the effect of external forces on a solid object.*

*The other group of goals related to changes in cognitive behavior is more general in nature and reach beyond the course. Some examples are to increase the students' ability to:*

- *Understand the basic concepts in statics,*
- *Solve out-of-context problems,*
- *Visualize 3-D problems,*
- *Communicate technical material orally.*

*Note that some of the goals in each group focus on technical issues, like problem solving, while others concern more general or professional skills, such as visualization and communication.*

*Goals in the second category, that is goals related to affective behavior, focus on how the students see themselves and the world around them. Some examples are to improve the students':*

- *Interest in the course material*
- *Attitude about the profession, the curriculum, or the department,*
- *Self-confidence,*
- *Intellectual development.*

*Goals for the next category concern changes in success rates. Some examples are to increase the students':*

- *Recruitment rates,*
- *Retention or persistence rates,*
- *Graduation rates.*

*The last category deals with diversity issues and basically looks at goals in the other three categories as they relate to some subpopulation or targeted subgroup. These address NSF's emphasis on broadening the participation of underrepresented groups. Some examples are to increase a targeted group's:*

- *Understanding of concepts,*
- *Processing skills,*
- *Achievement rate,*
- *Attitude about profession,*
- *Confidence,*
- *Intellectual development.*

***Exercise #2 - Writing Outcomes:*** *Two possible goals for the project described in the abstract above are: (1) increase the students' understanding of the concepts in statics and (2) improve the students' attitude about engineering as a career. You should write one or more outcomes for each of these goals.*

***Response #2 –Writing Outcomes:*** *Two possible outcomes for the first goal on conceptual understanding are: (a) students will be better able to solve simple conceptual problems that do not require the use of formulas or calculations, and (b) students will be better able to solve out-of-context problems. Two possible outcomes for the second goal on attitude changes are: (a) students will be more likely to describe engineering as an exciting career, and (b) the percentage of students who transfer out of engineering after the statics course will decrease. Please remember that these are examples of possible outcomes and there are many others.*

***Exercise #3 – Writing Evaluation Questions:*** *Two possible measurable outcomes for the project described in the abstract above are: (1) students will be better able to solve simple conceptual problems that do not require the use of formulas or calculations, and (2) in informal discussions, students will be more likely to describe engineering as an exciting career. You should write one or more evaluation questions for each of these outcomes.*

***Response #3 – Writing Evaluation Questions:*** *There are two general forms these questions take: one asking if a change occurred and the other asking if the intervention caused a change to occur. For the conceptual understanding outcome, these questions become: (a) did the students' ability to solve simple conceptual problems increase and (b) did the use of the 3D rendering and animation software increase the students' ability to solve simple conceptual problems. For the attitude change outcome, the evaluation questions become: (a) did the students discussions indicate more excitement, about engineering as a career, and (b) did the use of the 3D rendering and animation software increase the students' excitement about engineering as a career in their informal discussions.*

## TOOLS FOR EVALUATING COGNITIVE OUTCOMES

After completing the initial step in the evaluation process, defining goals, outcomes, and evaluation questions, the project team must next select approaches for collecting data to answer the evaluation questions. This involves developing a protocol and selecting appropriate tools and methods. The project team must develop their protocol and, for example, identify the test group and the control or comparison group, if control comparison design is planned. They also need to specify the timing, conditions, and individuals responsible for each evaluation activity. These protocol issues are important and complex, and we do not have space to explore them here. However, this importance and complexity supports the need for evaluation expertise on the project team as it develops its protocols.

In conjunction with developing the protocols, the project team also needs to identify or develop appropriate tools. These tools take on many forms and the NSF handbook on project evaluation [1] and the review by Olds and her colleagues [2] describe some of them. Some of the common tools are surveys, interviews, focus groups, and observation. Each has several variations, for example, survey questions may involve selecting one of several predefined choices (i. e., forced choices) or they may require open-ended responses or a mix of both types. Similarly, interviews may be structured with a fixed set of questions or they may be free-flowing and in-depth with the sequence of questions guided by the

responses. Interviews and focus groups differ in that the latter allows for the interaction among the participants to affect the flow of ideas. The selection of the specific tool should be based on what you want to know about project outcomes and the intervention.

The NSF handbook on project evaluation [1] and the review by Olds et al [2] summarize the properties of several approaches and list advantages and disadvantages of each. For example, the review by Olds et al [2] indicates that surveys are efficient but difficult to develop, particularly ones with established validity and reliability. They also note that surveys can be difficult to interpret when the response rate is low and that the accuracy of the responses depends on the subject's honesty. In discussing observations, these authors note that this approach is useful for observable behavior, captures behavior that subjects are unlikely to report, is time and labor intensive, and, when multiple observers are involved, requires training and an investigation of inter-rater reliability.

As a second example of the type of information in these references, the NSF handbook [1] describes situations for which various tools are appropriate. For example, interviews are appropriate for these types of evaluation questions:

- What does the program look and feel like?
- What do stakeholders know about the project?
- What are stakeholders' and participants' expectations?
- What features are most salient?
- What changes do participants perceive in themselves?

As a specific example of an evaluation tool, we want to consider the concept inventory, a tool that measures conceptual understanding. A concept inventory is a series of multiple-choice questions in which each question focuses on a single concept. Selecting the best answer should not require any formulas, calculations, algorithms, or problem solving skills. The initial concept inventory, the Force Concept Inventory (FCI), focuses on the mechanics portion of physics, and its use has changed the way introductory physics is taught in many places [3], [4]. Various groups are developing concept inventories for a number of engineering-related areas, including chemistry, statistics, strength of materials, thermodynamics, heat transfer, fluid mechanics, circuits, signal and systems, and electromagnetic waves [5].

Developing a reliable concept inventory for a given area is complex and time-consuming. The developer must identify the difficult concepts, the misconceptions, and the detractors; develop, test, and refine a set of questions and answers; establish the validity and reliability of tool; and deal with all of the ambiguities and multiple interpretations inherent in our language.

***Exercise #4 – Tools for Cognitive Outcomes:*** *A team of faculty members is developing a new approach that they believe will increase the students' understanding of the basic concepts in their course. In the literature, they have found a concept inventory that addresses their area. Write a list of questions that they should consider as they determine whether or not they will use this tool in their project evaluation? Again, we urge you to approach this as an exercise and take a few minutes to write down your response before reading the program director's responses.*

***Response #4 – Tools for Cognitive Outcomes:*** *The program directors' responses included questions in three areas related to the nature, testing, and prior use of the tool. One group of questions addressed the nature of the tool and the match between the tool and their evaluation task. Examples of this type of question were:*

- *Is the tool relevant to what was taught?*
- *Is the tool competency based?*
- *Does the tool assess conceptual or procedural knowledge?*

*The next set of questions dealt with the prior validation of the tool and examples were:*

- *Has the tool been tested?*
- *Is there information on reliability and validity*
- *Is there data comparing the tool to others?*
- *Is the tool sensitive? Does it discriminate between a novice and an expert?*

*The final set of questions concerned the experience of others with the tool and examples were:*

- *Has the tool been used by others besides the developer? At other sites? With other populations?*
- *Is there normative data?*

## TOOLS FOR EVALUATING AFFECTIVE OUTCOMES

Concept inventories, the tool that we just considered, provide data for evaluating cognitive behavior changes and many faculty members are comfortable with the idea of measuring cognitive skills because it represents a natural extension from the familiar approaches used in testing for understanding in the grading process. However, tools that provide data on affective changes are less familiar and so a little more daunting. Various investigators have developed tools for measuring students' attitudes, perceptions, and beliefs, including tools for measuring:

- Perceptions about the profession, the department, working in teams, etc,
- Attitudes toward learning,
- Motivation for learning,
- Self-efficacy, self-confidence,
- Intellectual development,
- Ethical behavior.

***Exercise #5 – Tools for Affective Outcomes:*** *Suppose your project's outcomes included improving the students' perceptions about the profession and their intellectual development and you want to find a tool that has been vetted. (The word "vetted" refers to established, tested tools.) Indicate whether you believe that you will find vetted tools for each type of measurement and whether*

*you believe that some of those tools will be quantitative in nature.*

***Response #5 – Tools for Affective Outcomes**: As we will see in the next two paragraphs, vetted tools do exist and some are quantitative and some are qualitative.*

One example of a quantitative tool for measuring students' attitudes towards engineering and their self-efficacy is the Pittsburgh Freshman Engineering Survey [6]. This is a questionnaire that asks students about a number of issues, such as their confidence in their skills in chemistry, communications, engineering, etc and their impressions about engineering as a precise science, as a lucrative profession, etc. Each question has a set of predefined answers from which the students select the most appropriate response. Shuman and his colleagues [6] have validated this tool using item analysis, verbal protocol elicitation, and factor analysis. After validating this tool, they used it to compare the attitudes of students who stayed in engineering to those of students who left.

Felder and Brent [7] discuss tools for characterizing intellectual development, which characterizes how students see knowledge, beliefs, and authority. On one end of this spectrum, some individuals believe that "knowledge is absolute," th*at* there is only one correct answer, that the instructor is an authority, and that education involves learning the correct answer from the instructor. At the other extreme, some believe that "knowledge is contextual," that there are many answers to a question, that the instructor is an experienced consultant, and that education consists of learning how to analyze and select appropriate answers. The authors list the following approaches for measuring intellectual development:

- Measure of Intellectual Development (MID),
- Measure of Epistemological Reflection (MER),
- Learning Environment Preferences (LEP).

Looking a little broader, Turns et al [8] discuss qualitative and quantitative approaches for characterizing students' skills, attitudes, and characteristics in the following areas:

- Communication capabilities,
- Ability to engage in design activities,
- Perception of engineering,
- Beliefs about abilities,
- Intellectual development,
- Learning Styles.

## ANALYSIS AND INTERPRETATION OF EVALUATION DATA

Once evaluation data are collected using the selected protocols and tools, they must be analyzed and interpreted. Typically these types of measurements exhibit relatively large variability and modest sensitivity, and they are corrupted by large extrinsic factors that cannot always be controlled. Statistical approaches are routinely used to test for significance or to establish confidence measures.

***Exercise #6 – Interpreting Evaluation Data:** Table. 1 shows hypothetical concept inventory data, where each question in the inventory has four possible answers. Each row in the table corresponds to a question on the inventory (i. e., a concept being tested); the last two columns show the percentage of the students who selected the correct answer before ("Pre") and after ("Post") some intervention or change in instruction was implemented in a course. Consider the percentages for Concepts #1, #2, and #3 and select the best answer for the following statements for each question:*

1. *The concept tested by the question was: (a) easy, (b) difficult, or (c) can't tel.*
2. *Understanding of the concept tested by the question: (a) decreased, (b) increased, or (c) can't tell.*

**TABLE 1.**
**HYPOTHETICAL CONCEPT INVENTORY DATA**

| Question (Concept) | No. of Students | | Percent with Correct Answer | |
|---|---|---|---|---|
| | Pre | Post | Pre | Post |
| 1 | 25 | 30 | 29 % | 23 % |
| 2 | 24 | 32 | 34 % | 65 % |
| 3 | 25 | 31 | 74 % | 85 % |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |

***Response #6 – Interpreting Evaluation Data:** Concept inventories provide no information about the difficulty of the concept and so the best answer for Statement # 1 is (c) "can't tell" for all three concepts. Although the percentage for Concept #1 decreased after the intervention, the change is small and both values are near 25 %, which is the expected percentage for random selection (i. e, based on a random selection of one out of four choices). For Concept #2 the change is large enough to indicate that there probably was some increase in understanding. The small difference in Concept #3 is probably within the variability of the measurement and so it is unlikely that there was a change. However if the number of students were large, this could be a statistically significant difference. As noted earlier, there are statistical tests that can be used to determine whether changes in data like these reach statistical significance or to determine a measure of the confidence that the change is real.*

Once a project team decides that the evaluation data indicates a change, it has a second and much more difficult task in trying to attribute the change to the intervention. Recall that when we discussed evaluation questions in Exercise #3, we noted the two types of questions, which basically asked: "was there a change" and "can the change be attributed to the intervention". There are many factors that make attribution difficult and a project team should try to

identify as many of these as they can ahead of time and then design their evaluation protocol to minimize the effect of these confounding factors, as they are commonly called.

***Exercise #7 - Alternate Explanation:*** *The data for Concept #2 in Table 1 suggests that there was a change in understanding. One interpretation is that the intervention caused the change but there are confounding factors. List some alternative explanations for the observed change.*

***Response #7 - Alternate Explanation:*** *Possible alternate explanations for the increased level of understanding are:*

- *Students learned the concept through some other mechanism (e. g., in another course or in study groups with students not in the course),*
- *Students in Post test gave answers that they thought the instructor wanted rather than what they believed or "knew",*
- *An external event (big test in previous period or a "bad-hair day") distorted pretest data,*
- *Instrument was unreliable,*
- *Other changes in the course and not the intervention caused improvement,*
- *One or both student groups were not representative.*

***Exercise #8 – Masking Factors:*** *The data for Concept #3 in Table 1 suggests that there was no change in the understanding. It is possible that there really was a change in the level of understanding but other factors (i. e., confounding factors) masked the change. List possible confounding factors that could have masked a real change resulting from the intervention.*

***Response #8 – Masking Factors****: Possible confounding factors are:*

- *An external event (big test in previous period or a "bad-hair day") distorted post-test data,*
- *The instrument was unreliable,*
- *Implementation of the intervention was poor,*
- *Population was too small,*
- *One or both student groups were not representative,*
- *Formats were different on pre and post tests.*

Typically, an interpretation of what a student knows or does not know would not be based on a student's answer to one item or one piece of data. Additional pieces of information are used when making such a judgment. This logic is similar for evaluating and intervention to assess if the intervention made a difference. Multiple types of data should be used when making a judgment about the effect of an intervention.

In closing this section we want to point out that cultural differences affect evaluation results. Because of this, evaluations should be done with awareness of the cultural context of the project. Evaluations should be responsive to racial and ethnic diversity, gender, disabilities, and language.

## DEVELOPING AN EVALUATION PLAN

Since we have considered several aspects of an evaluation process including the definition of the goals, outcomes, and evaluation questions, the selection of the protocol and the tools and the analysis and interpretation of evaluation data, we now turn our attention to the development of an evaluation plan. Before we begin working on the evaluation plan, it is important to understand the limitations of a project evaluation. A successful evaluation can provide reasonably reliable and reasonably valid information about the merits and results of a particular program or project operating in particular circumstances. It is not going to provide indisputable, objective evidence that an approach will always and everywhere provide the benefits seen in the evaluated environment. Generalizations are always tenuous.

***Exercise #9 – Evaluation Plan Topics:*** *The goals of a proposed project are to increase the students' understanding of the concepts in statics and to improve their attitude about engineering as a career. List the topics that should be included in the project's evaluation plan. List these topics in the order you would address them in a written evaluation plan so that they form a rough outline of the plan.*

***Response #9– Evaluation Plan Topics****: Program directors developed the following list of topics*

- *Name and qualifications of the evaluation expert,*
- *List of goals and outcomes and evaluation questions,*
- *Tools and protocols for evaluating each outcome,*
- *Procedure for analysis and interpretation of results,*
- *Possible confounding factors and approaches for minimizing their impact*
- *Formative evaluation techniques for monitoring and improving the project as it evolves,*
- *Summative evaluation techniques for characterizing the accomplishments of the completed project.*

*This sequence represents one of many possible choices and you should not infer that this is the "correct NSF approach."*

At this point, you should realize that evaluation is a complicated process and that the successful design and implementation of this process requires real expertise and cannot be an "add-on" at the end of the project. We recommend that a project team include an evaluator from the start of the project planning effort. Besides leading to a sounder evaluation effort, the evaluator's unique focus and perspective will shape the way the project evolves and lead to a more carefully defined project than one done without evaluation expertise integrated from the start. Certainly, we recommend that all education proposal writing efforts involve an evaluator, not only because this will lead to a more compelling evaluation plan than one done by typical STEM faculty members, but also because this perspective will lead to a better project.

A project team can look to the following to find an evaluation collaborator:

- Other departments (e. g., departments of education, educational psychology, psychology, administration, sociology, anthropology, science or mathematics education, engineering education),
- Campus teaching and learning centers,
- Colleagues and researchers,
- Professional organizations,
- Independent consultants,
- NSF workshops or projects.

Many times the evaluation effort relies on established techniques and, since it does not represent innovative work within the field of evaluation, it becomes somewhat routine and the evaluator functions in a "service capacity". This type of role is fine for professional evaluators whose career depends on successful involvement in evaluation projects. However, for tenure-track faculty members, who need to establish a record of research accomplishments in their discipline, the service role can be a problem. With some creative thinking, it is possible to define the evaluation process in a way that will serve the needs of the project and, at the same time, offer real research opportunities for the evaluator.

***Exercise #10 - Evaluator Questions:*** *Consider the perspective of an evaluator who has been invited to join a project team working on an education project. Write a few questions that this evaluator might have about the project and his or her involvement.*

***Response #10 - Evaluator Questions:*** *Questions fit into two categories – those related to project issues and those related to operational issues. Examples questions related to project issues are:*

- *What are the goals and the expected measurable outcomes*
- *What are the purposes of the evaluation?*
- *What do you want to know about the project?*
- *What is known about similar projects?*
- *Who is the audience for the evaluation?*
- *What can we add to the knowledge base?*

*Example questions related to operational issues are:*

- *What are the resources?*
- *What is the schedule?*
- *Who is responsible for what?*
- *Who has final say on evaluation details?*
- *Who owns the data?*
- *How will we work together?*
- *What are the benefits for each party?*
- *How do we end the relationship?*

## WORKING WITH AN EVALUATOR

As a STEM faculty member begins to think about working with an evaluator, there are some preliminary activities that will facilitate the start of an effective working relationship. First, STEM faculty members should recall their previous experiences with evaluators and then they should ask their colleagues about their experiences in terms of what worked and what did not work. They should clarify the purpose of their project and the evaluation, particularly focusing on the project's goals and outcomes, on the evaluation questions, and finally on the anticipated usefulness of the evaluation. Finally they should try to identify all the confounding factors they anticipate and the alternative explanations for the outcomes of the project.

As you begin working with an evaluator, it is important to develop a shared vision for the project, particularly for the evaluation effort. The project team, which includes the evaluator, should make sure that the project has clear goals, objectives, outcomes, and evaluation questions and that the project's planned activities are related to them. They should anticipate results and consider possible unanticipated positive outcomes and negative consequences. The team should discuss strategies for dealing with negative findings. As a way to increase the success of the project and the usefulness of the evaluation, the team should develop a project logic model to guide their efforts. As team members work together over time, they should begin to develop a team-orientation and ways to assess the effectiveness of their relationships on a continuous basis.

## SUMMARY

In this paper, we have tried to help you develop a better understanding of the evaluation process and the tools and techniques used in such a process to enable you to work with an expert evaluator in preparing and implementing an evaluation plan for an education development project. We have focused on the role and importance of project goals and outcomes in the evaluation process, on the nature of various types of outcomes, on evaluation tools for monitoring these types of outcomes, and on the issues involved in the interpretation of evaluation data. We suggested several issues to consider when writing an evaluation plan for a proposed project and in working with an evaluator.

## ACKNOWLEDGEMENTS

We would like to acknowledge the contribution of other program directors, particularly Patrick Carriere, Ken Gentili, and Roger Seals, in developing earlier versions of the workshop material.

## REFERENCES

[1] The 2002 User-Friendly Handbook for Project Evaluation, National Science Foundation, NSF 02-057, 2002.

[2] Olds, Barbara, M., Barbara M. Moskal, and Ronald L. Miller, "Assessment in Engineering Education: Evolution, Approaches, and Future Collaborations", *J. Eng. Ed*., Vol. 94, No. 1, 2005, pp. 13-25.

[3] Hestenes, D., M. Wells, and G. Swackhamer, "Force Concept Inventory", *The Physics Teacher*, Vol. 30, 1992, pp. 141-158.

[4] Mazur, Eric, "Qualitative versus Quantitative Thinking: Are We Teaching the Right Things? "*Optics and Photonics News*, Vol. 3, 1992, p. 38.

[5] Richardson, Jim, "Concept Inventories: Tools for Uncovering STEM Students' Misconceptions", in *Inventions and Impact,* AAAS, 2004, pp. 19-25.
[6] Shuman, Larry, J., Mary Besterfield-Sacre, and Jack McGourty, "The ABET 'Professional Skills' – Can They Be Assessed?" *J. Eng. Ed*., Vol. 94, No. 1, 2005, pp. 41-55.
[7] Felder, Richard M., Rebecca Brent, "Understanding Student Differences", *J. Eng. Ed*. Vol. 94,No. 1, 2005, pp. 57-72.
[8] Turns, Jennifer, Cynthia J. Atman, Robin S. Adams, and Theresa Barker, "Research on Engineering Student Knowing: Trends and Opportunities, *J. Eng. Ed*. Vol. 94, No. 1, 2005, pp. 27-40.