A Phase II Trial of the Systems Evaluation Protocol for Assessing and Improving STEM Education Evaluation

Year 1 Report of the External Advisory Board

This report summarizes the conclusions and recommendations of the External Advisory Board (EAB) for the Phase II Trial of the Systems Evaluation Protocol (SEP) for Assessing & Improving STEM Education Evaluation (NSF award 0814364). The report has four sections. The first outlines the EAB’s assessment of the project team’s progress in the first year. The second section describes the major issues and challenges for the project. A brief discussion of the potential applications of the project is provided in the third section. The report ends with a set of recommendations for the project team to consider as it moves forward.

**Assessment of Progress and Findings to Date**

The EAB was unanimous in its praise for the productivity of the project. The quantity and quality of the work as well as the teamwork demonstrated by participants are impressive for a project in its first year. This overall assessment is supported by our review of progress toward each of the project’s six goals.

1) **Assess the degree to which the SEP is associated with changes in organizational evaluation capacity and performance outcomes.**

   The measures used to assess changes in organizational evaluation capacity and performance outcomes have been identified/developed and baseline data collected for all programs. (However, see the EAB’s questions about the assumptions underlying the link of SEP to organizational capacity in the next section.)

2) **Assess the performance of alternative versions of implementing the SEP.**

   Information on the outputs of the traditional (face-to-face facilitation) SEP is being collected. Comparison to the virtual SEP will not be possible until the virtual version is complete.

3) **Assess the relative performance of the SEP in two major and different STEM education contexts.**

   The team has made good progress toward this goal. They have developed strong partnerships with two very different organizations (Cornell Cooperative Extension and Materials Research Science Education Centers) and have begun the implementation of SEP in its traditional form with both.

4) **Create a self-guided Virtual SEP for subsequent potential national testing and distribution.**

   Further development of the Virtual SEP is pending results of the formative evaluation of SEP.

5) **Enhance evaluation cyber infrastructure (the Netway) through the development and testing of researcher and policy-maker portals.**
The release of the new version of Netway in early 2009 and the monitoring of usage rates indicate strong progress toward this goal.

6) **Train the next generation of STEM education evaluators and enhance evaluation capacity within STEM education generally.**

The project has assembled an interdisciplinary team of Cornell undergraduate students, graduate students from two universities (including a student in the American Evaluation Association Diversity Internship Program), and STEM professionals. In addition, the team has been able to develop partnerships with more Cornell Cooperative Extension programs than originally anticipated. In combination with publications and presentations, these accomplishments indicate strong potential for training new evaluators and developing new evaluation skills in current STEM educators.

### Major Issues or Challenges

The EAB identified some key questions and considerations for the SEP project team. This section includes some specific recommendations in addition to those provided in the final section.

**How is systems thinking manifested in SEP tools and outputs?**

The EAB raised the question of how systems thinking is expected to translate into improved evaluation planning processes and plans. The anticipated linkages between the introduction of systems thinking in the SEP Facilitator’s Guide and the quality of the logic models and evaluation plans that result from using the SEP are not clear. The EAB also questioned the value added by systems approaches as described in the Facilitator’s Guide compared to other evaluation models. For example, the Facilitator’s Guide includes a lifecycle analysis as a reflection of “an evolutionary systems evaluation perspective,” but other evaluation frameworks that are not systems-oriented include the concept of program development and maturation. Similarly, logic models are a commonly used tool. What is it about the systems approach that make the lifecycle analysis or logic models better than what is achieved with other approaches?

In response to these questions, the project team suggested that SEP is likely to change the evaluation planning process. Specifically, the people responsible for developing the plan should experience increased connections to others doing similar work, have more sensitivity to the program context, and be more engaged in the evaluation. Speaking specifically to the question about program life cycle, the team pointed out the difference between maturing (getting older) and evolving (changing form in response to a dynamic system). While understanding the clarifications offered by the project team, the EAB recommends further explication of how systems thinking is manifested in the SEP, how it is an improvement over other approaches, and how it affects the quality of evaluation planning products (e.g., logic models, plans).
Will the SEP tools create a mechanistic approach to evaluation?

In part because of the lack of clarity about how systems thinking is manifested in the SEP implementation and outputs, the EAB raised the concern about SEP encouraging a mechanized approach to evaluation. In short, does the SEP substitute for judgment or does it facilitate judgment? For example, the worksheets in the Facilitator’s Guide can be used without thinking about underlying systems issues. Netway, despite its exciting implications, seems to have particular potential for supplanting evaluator judgment with a sort of “paint by numbers” approach: Program staff can go to Netway, identify evaluation components that have already been developed, and apply them without considering the values and needs of their program environment and stakeholders. The EAB acknowledged that tools can lead to new ways of thinking, sometimes in ways that are not recognized by the person using the tool. They nonetheless urged the team to be aware of the potential negative effects, such as a lack of engagement and judgment in evaluation planning and implementation. The team indicated that they are aware of this concern and are planning to include critical reflection prompts in the virtual SEP. They also suggested that this is the kind of effect that might be revealed by the comparison between the traditional and virtual versions of the SEP.

What is the link between individual skills and organizational capacity?

Going beyond SEP outputs to consider its intended effects on evaluation capacity, the EAB identified two issues. First, the placement of increased evaluation capacity at the end of the planning process seems premature. The EAB recommended revising the SEP path model so that changes in capacity come after evaluation implementation.

The second issue associated with evaluation capacity as framed by the project team is the apparent assumption that intervening at the level of the individual will change organizational capacity for evaluation. The EAB asked if the team has considered intervening at the organizational level, possibly influencing evaluation policy, in addition to building skills of individuals. The team responded that at this point they are focused on the development and testing of SEP; the context issues are beyond the scope of the project. The EAB pushed for greater clarity about the degree to which organizational capacity is the goal and, to the extent it is, for a stronger argument for the link between the development of individual skills and ways of thinking about one’s program and change in organizational capacity.

While discussing the theory of change depicted in the path model, the EAB also noted that project materials indicate ongoing relationships between the project team and project partners. This seems like an important aspect of project implementation that is not currently and should be captured in the path model. The EAB also encouraged the team to think about how their role in implementing the traditional version of SEP will be duplicated in the virtual version.
How will the intended comparisons work given the likelihood that the tools will be adopted across groups?

The project has already experienced considerable success in generating interest in the tools – both the SEP Facilitator’s Guide and Netway. Given that success, the EAB questioned how the team will deal with contamination of potential comparison groups. The team responded that they think of it more positively as “contagion”, a catching on of the approach. As they see it, their task is not to discourage adoption but to make sure they track the nature and extent of dissemination across cohorts. The EAB nonetheless indicated the need to consider how the challenges in making comparisons will affect their ability to answer their evaluation questions.

Implications and Applications of Project Work

The EAB identified four important implications and applications of the project over time:

• The project has an opportunity to contribute to the body of knowledge about evaluation capacity building by examining the link between individual skills and organizational capacity.

• The cyber infrastructure of the project has the potential to improve the quality of STEM evaluation by facilitating the sharing of materials and providing a structure for learning across types of STEM programs and across sites using similar interventions. A related implication is the opportunity for SEP users to increase their connections with fellow educators and evaluators.

• Netway is a powerful tool that could facilitate evaluation planning and implementation in a wide range of contexts, within and beyond STEM evaluation.

• The students and partners working with the SEP project are gaining a variety of skills. In addition to evaluation-related skills, they are also gaining experience with technological aids to evaluation planning and learning how to work across disciplines.

Consultative Advice and Recommendations

As stated at the beginning, the EAB is impressed with the progress of the SEP project team. We have provided some recommendations related to specific challenges in an earlier section. We offer the following additional advice in hopes of strengthening an already outstanding project.

Defining the boundaries of the project

In the longer-term, the EAB advises the team to think about bounding its ambitions – is the goal changing evaluation capacity at a macro level? Or is the focus really on individual and local levels? We applaud the bottom-up approach, but suggest that the team expect very modest effects at the higher levels of the program/organizational system.

While we recommend that the team limit its ambitions for capacity building at the lower levels, we also encourage them to think about how they could intervene at the more macro levels. For example,
in what ways can they communicate the value added of a non-mechanized use of the tools to organizational leaders?

**Clarifying evaluation plans**

Three issues associated with the evaluation plans came up – maintaining clear comparisons, addressing the potential conflict between the development of Netway for commercial purposes and the evaluation of Netway’s role in SEP, and developing criteria for assessing progress toward the goal of training the next generation of STEM education evaluators. We advise the team to attend to these issues in the next phases of the project.

**Conceptualizing and operationalizing systems thinking**

We recommend further reflection on the effects of systems thinking on the SEP users and their evaluations. Beyond the conceptual framework of the systems approach and the pedagogical value it offers, what are the concrete contributions of systems thinking, as operationalized in SEP and Netway, to evaluation planning and implementation? In what ways will system thinking manifest itself? At what points in the evaluation cycle is it likely to appear?

If the effects are not likely to be observed until the implementation stage, then it would be valuable to begin now to think about how the differences from non-SEP implementation will be evaluated. As an immediate step, the scoring rubrics for the logic models and plans should be reviewed to see if there are aspects of the systems approach to evaluation that should be being tracked.

**Conclusion**

In its first year, the SEP project developed numerous partnerships and produced an initial set of tools. The project seems to be on track to achieve the research goals. The EAB has some concerns about the behind-the-scenes role of systems evaluation in the project and encourages the team to make the systems focus of the project more explicit in all aspects of the work. In the short-term, the EAB sees the value of SEP products, including Netway and the Facilitator’s Guide, as facilitators of evaluation planning. In the long-term, there is potential for the project to positively affect the quality of STEM evaluation more broadly.

*Submitted October 5, 2009*

**External Advisory Board Members**

Leslie J. Cooksy, Chair, University of Delaware  
Jennifer C. Greene, University of Illinois  
Rodney K. Hopson, Duquesne University  
Melvin M. Mark, Pennsylvania State University  
Debra Rog, Westat