ANNUAL REPORT FOR AWARD # 0814364

CornellUSt:Aawds PRIOR 5\10
A Phase II trial of the Systems Evaluation Protocol for Assessing and Improving STEM Education Evaluation

Participant Individuals:
Senior personnel(s) : Jane C Earle; Claire E Hebbard
Graduate student(s) : Tom G Archibald
Undergraduate student(s) : Haixin Dang; Stacey P Eilbaum; Nicole Monahan; Joshua Hatcher-Mullins
Other -- specify(s) : Leslie Cooksey; Mel Mark; Jennifer Greene; Rodney Hopson; Debra Rog
Senior personnel(s) : Monica Hargraves
Graduate student(s) : Margaret Johnson
Other -- specify(s) : Jennifer Urban
Graduate student(s) : Wanda Casillas
Technician, programmer(s) : Jamila W Simon
Graduate student(s) : Larry Daffin
Undergraduate student(s) : Matthew Negrin; Drew Stadlen
Graduate student(s) : David Carucci; Linda McMahan
Undergraduate student(s) : Natalie E Cook; Kelsey N Brykman
Other -- specify(s) : Nevjinder K Singhota
Undergraduate student(s) : Elizabeth B McCoy; Miranda Fang
Graduate student(s) : Marissa Burgermaster
Undergraduate student(s) : Darlene Kim; Stephanie Phelan

Partner Organizations:
Cornell Center for Materials Research: In-kind Support; Facilities; Collaborative Research

For years 1-2 of this project, one of our .5 FTE key staff (Jane Earle, Nevjinder Singhota) is an employee at CCMR. She uses their office space, telephone and computer for her work with this project. Also, we maintain communication with CCMR leadership as we continue to discuss potential future collaborations.

Cornell Cooperative Extension: Financial Support; In-kind Support; Facilities; Collaborative Research; Personnel Exchanges

Many of our participants are CCE programs. Additionally, Cornell University Cooperative Extension funds the full-time position of the Manager of Evaluation for Extension and Outreach, and this person (Monica Hargraves, who is listed as working on the project) serves as the evaluation facilitator for the CCE cohort of this project. CCE and state 4-H leadership consults with us for selecting Science, Engineering and Technology (SET) programs, and Agricultural Programs to participate in our research. And the CCE
system, as a whole, benefits from the training their staff receive during our project.

Cornell University, College of Human Eco: Financial Support; Facilities

The College of Human Ecology at Cornell University sponsors a graduate research assistant (Margaret Johnson) who has participated the first two years of the project.

Montclair State University: Collaborative Research

Dr. Urban - project CoPI - is at Montclair State University.

Dr. Jennifer Urban, from MSU, is the Co-PI on this grant. She has been integrally involved in vSEP development and in the evaluation of this research.

Other collaborators:

Through CCMR we have connected with the national MRSEC group. Thirteen individual MRSECs are currently participating in our research, while also using our materials to improve their programming. Participating MRSECs include Yale CRISP, University of Chicago MRSEC, University of Maryland MRSEC, USC-B MRSEC, Penn State MRSEC, University of Washington MRSEC, University of Nebraska MRSEC, University of Minnesota MRSEC, University of Pennsylvania MRSEC, University of Mass, Amherst, MRSEC, University of Ohio MRSEC, Northwestern University MRSEC, Colorado School of Mines.

Similar to these three MRSEC programs, there are 46 CCE programs that also are working on this project. This includes the leadership and 4-H program staff from the following counties (or region) in New York State: Chemung, Chenango, Clinton, Cortland, Franklin, Fulton, Genesee, Jefferson, Lake Plains (Monroe and Wayne counties combined), Madison, Oneida, Ontario, Oswego, Rensselaer, Saratoga, Seneca, Tioga, Tompkins, and Ulster, pilot project sites from Chemung, Tompkins, St. Lawrence, Ulster and Jefferson counties, and the two Regional Vegetable teams at Cornell University.

Future collaborations are likely to increase because there has been a wide-spread interest in the project. The MRSECs are showing a greater interest in getting involved than has been supported through our current funding.

Activities and findings:

Research and Education Activities:

This report is for Year 3 (August 2010 – July 2011). This research grant constitutes one phase of a four-phase model for developing and testing the effectiveness of the Systems Evaluation Protocol (SEP)
in evaluating STEM outreach programs. This study is a Phase II trial. As such it is not meant to be a causal assessment of the effects of SEP; it is meant to assess whether the SEP program, when implemented in several different common STEM contexts, is associated with change in program evaluation capacity and quality when evaluating short and intermediate outcomes. There were three distinct research subprojects to be conducted throughout the 5-year research plan. These included 1) the phase II study of the facilitated SEP to assess if it is associated with change in program evaluation capacity; 2) the development of a virtual SEP (vSEP) and comparison of this method to the facilitated method; and 3) the continued development and testing of the Netway - an evaluation cyberinfrastructure that facilitates networking, program modeling, and evaluation planning.

There are two non-research components of this project, 1) an External Advisory Board, and 2) sponsoring and hosting a Graduate Diversity Intern with the American Evaluation Association. This year we also received a supplement for undergraduate mentoring.

Major research activities and education activities (experiments, observations, presentations):
1. The facilitated SEP.

The general process of our research is to do use the Systems Evaluation Protocol (SEP) to guide STEM outreach programs on evaluation planning, and then to assess the quality of their plans. We also assess evaluation capacity of both the program and its home organization to look for change over time. Programs are recruited to go through Evaluation Planning training in Treatment Year 1, and if that is completed they are given the option to enroll for Evaluation Implementation support in Treatment Year 2. A new cohort begins each year, and ideally proceeds as outlined in the chart below (Cohort 1 was in the previous project, so this project begins with Cohort 2).

<table>
<thead>
<tr>
<th>Phase II trial of the Systems Evaluation Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1            Y2           Y3       Y4Y5</td>
</tr>
<tr>
<td>Cohort 2 Treatment1* Treatment2**</td>
</tr>
<tr>
<td>Cohort 3 Treatment1 Treatment2</td>
</tr>
<tr>
<td>Cohort 4 Treatment1 Treatment2</td>
</tr>
<tr>
<td>Cohort 5 Treatment1Treatment2</td>
</tr>
</tbody>
</table>

*Treatment1 = Facilitated Evaluation Planning Training base on the SEP, and **Treatment2 = Facilitated Evaluation Implementation Training/Support.

The primary focus of Year 3 has been to wrap up the facilitated Systems Evaluation Protocol on evaluation planning (fSEP Treatment1) with our 3rd cohort, then to start them on fSEP Treatment 2, while simultaneously Cohort 4 on fSEP Treatment 1). There is some overlap between project years and treatments, and therefore during this year Cohort 2 actually finished up their work with us on their fSEP Treatment2. They continue to stay in touch with each other and us on continuing collaborations and discussions.

This year's fSEP Treatment 1 (Cohort 4) consisted of 10 programs (5 MRSECs and 5 Cooperative Extension Energy team programs). The fSEP Treatment2 (Cohort 3) consists of 3 MRSEC programs and 4 Extension vegetable programs. This was down from the corresponding fSEP Treatment1 which was 10 programs; 3 programs from Cohort 3 evaluation planning did not continue on with implementation work due to staff turnover or time constraints.

The CCEs and MRSECs are considered two different systems, but both are doing science outreach. The CCE system was facilitated by the CCE funded Manager of Evaluation for Extension and Outreach, Monica Hargraves. The MRSEC system was facilitated by the research assistant funded through this grant - Jane Earle Buckley. Their work was supported by all other active project staff and students.

The Systems Evaluation Protocol (SEP) is the underlying protocol (step-by-step
process) that the treatments are based on. As we have learned what is important and how the concepts are best understood, the SEP, and the accompanying guide, has undergone revisions. The Guide for SEP Evaluation Planning is now finalizing the 2nd revision, whereas the SEP Guide for Evaluation Implementation and Utilization is undergoing its first version. One of the issues our External Advisory Board (EAB) had concern about what we meant by 'Systems Evaluation'. We worked this year to articulate more in-depth on the systems thinking around this approach and how it is new and different than traditional evaluation.

2. **vSEP**

The secondary focus of Year 3 was to finalize decisions on the vSEP (virtual SEP). After putting out RFPs for development of the vSEP we realized that comprehensive portal development was beyond the ability of this project (costs ranging approximately from $100,000-$600,000). The decision was made to develop an 'enhanced Netway' rather than the vSEP. The vSEP would have been a stand-alone site, much like an online course for evaluation planning, and would have replaced our facilitation process (which uses the Netway cyberinfrastructure as a tool for planning). The Enhanced Netway instead allows us to integrate the SEP training materials and recordings into our Netway (the proof-of-concept cyberinfrastructure that is funded on this project.) This was in part a financial decision, but was also based on discussions among ourselves and our EAB. Another concern of the EAB was that the vSEP would be too mechanistic of an approach to evaluation. Evaluation planning requires evaluative thinking skills, and trying to create a vSEP that would do justice to the complexity of the facilitation process of our training workshops (in order to develop the evaluative thinking) was financially beyond the scope of this project. Putting resources into a vSEP that we knew would not be able to compare to our intentions around using the SEP for evaluation planning would have been a waste of resources. Therefore, to be fiscally responsible as well as to get the most benefit from the Netway we decided to integrate our training and other resource materials into the Netway for reference to participants working on their evaluation plan. Rather than compare the fSEP to the vSEP, we will plan to compare the fSEP with Netway access to our Cohort #5, which will not be facilitated, but will have access to all the recordings and materials of previous cohorts. This will, in essence, be similar to the vSEP with much lower costs - since the materials are already developed for this method and would not require additional contracting with website developers. The revisions to the Guide to Systems Evaluation Protocol, mentioned above, has been rewritten with the intended use to be for self-facilitation of program staff doing evaluation planning (and implementation). (The previous guide was written for evaluators to use in facilitating others developing evaluation plans.)

3. **The Netway**

Netway improvements this year were structural to build the foundation for the Enhanced Netway format that will be replacing the original vSEP plan. We worked closely with our Netway developers to create a better Netway and a step towards social networking:

- A new and improved user interface.
- A much faster and easier Netway to work with.
- Logic modeling and Pathway models that are more closely integrated with our training, and user intuitive.
- Evaluation planning worksheets built into the Netway to encourage the logic and evaluative thinking process and to assist in writing the evaluation plan.
- The use of play lists that to streamline communication with similar programs.
- A new and improved Google-type search that is faster and much easier to use.
- WYSIWYG editing and template reports that are better tailored to reporting
needs.

• An improved measures database that will allow users to contribute their own measures.

4. **External Advisory Board**

The EAB met in person in July 2010, and via teleconference in January and May. Year 3 meeting is scheduled for August 5, 2011. Their advice has challenged us to really consider the scope of our project, encouraged us in how to direct our research efforts, and suggested we clarify our meaning of Systems Evaluation. Our research still focuses on evaluation planning, but the support requested by participants for evaluation goes beyond planning. They need help conducting and analyzing their data. Without EAB encouragement to stay focused it would be simple to quickly expand activities into other areas of reported need. Based on where this project stood a year ago, we are much more stable, having fewer changes to our processes, and are moving toward a potential Phase II study of the SEP.

Secondly, the EAB is encouraging more focus on the qualitative data abounding in this project, and perhaps using a mixed methods approach for analysis. Finally, their lack of clear understanding what we mean by Systems Evaluation has prompted us to better articulate the theories behind the SEP, and to interact with other Systems thinkers at the American Evaluation Association. Their Year 2 report is attached at the end of this report.

5. **AEA Diversity Intern.**

This year we did not have a graduate AEA Diversity intern. We had obtained permission instead to use those funds to retain the previous year's intern as a temporary research assistant. However, in late Fall he reported he was unable to commit time to the position. The plan is to attempt to have two interns next year. We are actively recruiting at our institutions, which is a particular challenge at Montclair State University since it is not traditionally known as a research institution and has not yet started the Ph.D. Program.

6. **Student Mentoring**

Student Mentoring –

Graduates: Both sites support a graduate research assistant. The Cornell student participates fully in the facilitation process of the project, working directly with our staff on project facilitation. This is his third year with the project. He presents his own work at several national conferences, which this project in part supported. Throughout the past year he has been involved with many aspects of the project including:

1. the development and implementation of the facilitation of our partnerships (a. planning and facilitation of in-person meetings, webinars, and individualized consultations—including developing instructional materials, and b. incorporating pertinent scholarship on adult education and evaluation capacity building);

2. the progression of our research plan (a. refining and managing surveys, rubrics, interview protocols, and other measurement tools, and b. assisting with database design, data collection, data entry, and data management, and plans for data analysis);

3. The progression of the project’s theoretical aspects (a. successfully presenting and proposing to present at the American Evaluation Association conference on topics including evaluation quality, the measurement of evaluative thinking, and the role of evaluation in bridging the research-practice gap, and b. contributing the editing and writing of the revised Guide to the Systems Evaluation Protocol );

4. The mentoring of undergraduate research assistants (a. working alongside undergraduates on many of the above tasks, and b. collaborating with an undergraduate on our project's theoretical and practical use of mixed methods approaches);

5. Professional development (a. attending workshops and conferences on evaluation, qualitative inquiry, mixed methods, and adult education, b. incorporating his experience on this project into his dissertation study, and c. applying his experience on this project to other, related STEM evaluation contexts in a consultative role.
The graduate student at Montclair was new to the program in December 2010 (replacing a previous student). She has worked on data management and interpretation for the rubric reliability testing, rubric revision, qualitative analysis protocol and coding, and the measures code book and coding keys for the project's survey measures. She coordinates the undergraduate assistant staff at Montclair State, participates in organizational and planning meetings, and has presented work on rubric inter-rater reliability with Jennifer Urban at a Montclair State University College of Education and Human Services Faculty Scholarship Poster Session. Marissa is also conducting descriptive analyses for some of the basic survey measures.

Undergraduates: both sites were able to find and employ undergraduates. Each student is fully integrated into the work for each site, and each works with their local team to focus their activities.

At Cornell two undergraduate students who had worked full time the summer of 2010 attended the annual conference for the American Evaluation Association (AEA) in November, and assisted in analyzing data and developing presentation of this work at that venue. One student – a senior at the time, actually presented at the conference with us. Two more students joined the team in January and worked intensively on literature reviews, supporting the development of training materials, became familiar with the SEP and the facilitation process, and assisted in creating tutorials for the updated Netway. Over the summer Cornell students have attended trainings on Access databases, data analysis, and other professional opportunities offered through Cornell (literature reviews, Endnote, etc.).

Montclair has two undergraduates working on the program. They both started with the project in February 2011 and have worked on various tasks, including data entry, qualitative coding of interview data and writing the content of several Netway tutorials. They examined Netway data linkages and content in order to suggest Netway interface reports, and are creating the initial content for a Montclair State University webpage about the project. In addition to familiarizing themselves with the Systems Evaluation Protocol (SEP), the field of evaluations, and qualitative data analysis, they maintains notes on audit trail and data entry for the data generated at Montclair State.

Here is a summary from one of our undergraduates: 'As an undergraduate research assistant in Cornell Office for Research on Evaluation (CORE), I have learned a great deal about what evaluation is, why it is important, how evaluation happens, the various possibilities for evaluation, and what the outcomes of evaluation can be. I have learned about evaluation through searching for and reading academic articles, participating in and helping to facilitate evaluation trainings with partners, and attending and preparing to present at the American Evaluation Association conference. In particular, I have been studying Mixed Methods evaluation. As I have been learning about the strengths and benefits of incorporating qualitative and quantitative methods and integrating qualitative and quantitative data, I have classified CORE's project as an integrated mixed methods evaluation project, which employs qualitative and quantitative research methods and unifies the emerging data even before analyzing them. As I have been learning about evaluation, my insights have really influenced my interests and even shaped my career goals. Prior to working at CORE, I knew I wanted to pursue graduate studies in Human Development, but I was not sure about what I actually wanted to do with the degree. Now that I have been exposed to evaluation, I am sincerely convinced that evaluation is a crucial part of every domain I had ever seen myself making contributions to. Whether I work in scientific research, social work, therapy, counseling, or any other field in which work is being done to achieve some outcome, evaluation will be a necessary component of my work. Working at CORE has helped me understand how I can unify my interests in human development research, social work, and evaluation in my future career goals. At this point, I aspire to do research that will allow me to implement evidence-based
youth development programs. I expect evaluation to be a key component of this work, as I pilot and implement programs and try to figure out if they are working and how to improve them so that they can make the most meaningful change in children's lives.

Findings:

We have no new findings associated with the facilitated SEP beyond those identified in Year 2. In summary - we know that there is a demand for assistance with conducting evaluation and analyzing data - that simple evaluation planning is insufficient; programs participating with us report an improved understanding of their program's logic and improved ability to talk about their program with funders; once staff have learned to use our tools for evaluation they often use it on other programs, and sometimes for reasons other than evaluation (i.e. program development); there are as many difference within systems as there are between systems; and that evaluation capacity requires efforts at all levels of the organization, including leadership. We are still collecting data as to the efficacy of our SEP. We continue to get people interested in working with us and gaining access to the tools.

One of the major deliverables of this project is the measures of Evaluation Plan Quality. This year we have worked on refining the rubrics that will be used to assess the quality of evaluation plans - these include the Logic Model Rubric, The Pathway Model Rubric, and the Evaluation Plan Rubric. Evaluation plans have several categories, among them are evaluation questions, specific sampling strategies, the identification or development of measures to assess the evaluation questions, an evaluation design, an analysis plan, a reporting plan, and an implementation plan and schedule. The form that was previously used to provide feedback to participants was adapted in order to use it as a summative tool of evaluation plan quality that would allow the team to quantify the quality of participating programs' Evaluation Plans. We have finished our third version of this measure. Revisions were based on some findings, such as: Need to promoting consistent responses from raters, need to eliminate ambiguity, need for reduction of repetition within and between subsections, if /then and nested responses were eliminated, Confirmation that additional knowledge of the program is not required beyond what is in the plan. Now on the third and we believe final version, we found inter-rater reliability for the Evaluation Plan Rubric to be 77%, of the Logic Model Rubric to be 93%, and of the Pathway Model Rubric to be 75%.

With the finalization of these measures we are ready to now begin completing rubrics on all the evaluation plans to date and will be able to begin preliminary analysis of this research project. However, data will continue to be collected throughout the next two project years.

Training and Development:

The evaluation facilitators on this project have continued to hone and deepen their skills and understanding of evaluation and teaching adult learners. Staff have participated at the national level in presenting their findings and in discussing issues of importance to professional evaluators. We continue to engage in discussions of systems evaluation theory.

Graduate students working on this project have been deeply engaged in
developing measures and conducting the research. They have had opportunities for teaching and public presentations (including being adept at PowerPoint), group collaborations, independent research, data tracking, and mentoring of undergraduate students.

Undergraduates working on this project are more confident in their abilities to pursue a graduate degree which includes research responsibilities. They have attended and participated in the face-to-face meetings with the Evaluation Partnerships, gained experience in several software programs - Word, Access, Powerpoint, Endnote, Illustrator - and the Netway cyberinfrastructure and in printing posters. Students are encouraged to take classes in webpage development, Access Databases, and Statistical Analysis (SAS, SPSS, Stata, Minitab, R).

**Outreach Activities:**

In November 2010 we presented our program (The SEP, the Evaluation Partnership process, and the Cyberinfrastructure) at the National Conference of the American Evaluation Association. The audience consisted of professionals interested in evaluation.

We have taken advantage of several opportunities to talk about evaluation of outreach programs with several audiences – both within New York State and beyond. These include statewide meetings of Executive Directors of the CCE Association, the National MRSEC Outreach Directors, and the National Cooperative Extension Evaluation forum. Due to this work and other work of the PI we are partnering with the newly developed Bronfenbrenner Translational Research Center at Cornell.

**Journal Publications:**


**Book(s) of other one-time publications(s):**

**Other Specific Products:**

**Software (or netware)**

Version VI of the Netway - the cyberinfrastructure for evaluation planning

This website and database - http://extensionnetway.net/ - is the cyberinfrastructure that is utilized by our evaluation partnership cohorts. and their colleagues. It is a password-protected resource so that most of the people using it are using the Systems Evaluation Protocol.
**Teaching aids**

We have developed measures and training materials associated with this project, and these are mentioned in the activities section but are also listed here. In addition to helping us to conduct our research these materials can be useful beyond our research.

These materials in their most current format will be made available in our next version of the printed protocol, and at that time will also be made available in the web.

**Internet Dissemination:**

http://www.core.human.cornell.edu

Website discussing research and disseminating research results

**Contributions:**

**Contributions within Discipline:**

The 'Facilitator's Guide to Systems Evaluation Protocol' was a significant contribution to the field of evaluation. We are using this publication with our current cohort, but also make it available both online (http://www.core.human.cornell.edu) and as a hard copy upon request. The new edition of this guide which had been drafted but not yet published will be a resource both within evaluation and in all other fields.

This grant is contributing to our general understanding of evaluation policy and its importance to evaluation practice and program improvement and development. This focus on how evaluation and evaluation policy affects all members of a system is key to creating an environment which supports evaluation.

The External Advisory board is composed of highly recognized evaluators. Their continued interest in the project contributes to the project as much as this project contributes to their own work.

The original measures for Organization Evaluation Capacity, Program Evaluation Capacity, Logic Model and Pathway Model Quality, and Evaluation Plan quality were released with the published Protocol, and the newer versions of these measures (See Findings for results on reliability testing of the Rubrics for Logic Model, Pathway Model and Evaluation Plan Quality).

**Contributions to Other Disciplines:**

The work of the cohorts in this project has improved their understanding of their program, as well as their ability to communicate with with their stakeholders about their goals and impacts. Additionally, these programs have a better understanding of evaluation, and are better equipped to conduct good evaluation of their programs.
As we have worked with programs to develop their evaluation plans the organization leadership has frequently followed the progress. In many cases the leadership is seeing their team's evaluation capacity and knowledge of the cyberinfrastructure as tools to help them in program planning and development.

Now that we have worked on developing a measure and running reliability on it we will be able to demonstrate this to programs in any discipline as an example of developing and testing new measures. This is extremely encouraging to our participants, who had the (false) belief that they could just take a pre-developed measure off the a shelf in order to evaluate their programs. The MRSEC system as a whole is now working together to find ways to collaborate on sharing and developing measures together, and trying to potentially do evaluation across the system.

**Contributions to Education and Human Resources:**

We are educating a new generation of practitioner-evaluators both in the NSF MRSECs and in CCE. In turn, these practitioners are spreading their knowledge to their peers, greatly enhancing the evaluation capacity of their organizations and of the STEM Outreach field as a whole. There has been an interest in offering a certification for the participants in this study who have demonstrated skill attainment and evaluation expertise. A certification would enhance their resumes. This is something that not every participant would achieve.

Opportunities for college students to work on this project have resulted in mentoring from CORE staff of several graduate and undergraduate students. Undergraduates particularly express an increased interest in pursuing scientific research as a career. They report that they have increased their communication with their science peers about science education and evaluation. Undergraduates will be taking leadership on writing and presenting an article for Mixed-methods database development and analysis at AEA in 2011.

**Contributions to Resources for Science and Technology:**

Updated versions of the Netway cyberinfrastructure, the Protocol, and the associated measures are all resources for programs and educators. Especially important, we believe, will be the new Guide to the Protocol (New in Year 3) that is written for outreach program staff at any level of experience in evaluation.

**Contributions Beyond Science and Engineering:**

Our research is contributing to the field of evaluation broadly by helping to shape the next generation of resources and processes that will be informed by the latest in systems thinking and approaches. We are contributing to the way the US federal government (and other entities) is approaching evaluation, primarily through our emphasis on evaluation policy (supported in part through this grant), but also through education on the idea of evolution of both programs and program evaluation. Many government agencies...
have required the use of the Randomized Controlled Trial (RCT) for program evaluation, without giving consideration to whether such a design is appropriate for the developmental level of the program. The SEP shows how RCT's for newer programs that are still in development may not be the most appropriate use of resources, and may not provide the evaluation data and rapid feedback required by these early-lifecycle programs.

The Systems Evaluation Protocol and cyberinfrastructure will be valuable to any field that engages in evaluation and program development.

Conference Proceedings:

Special Requirements for Annual Project Report:

Categories for which nothing is reported:
Conference Proceedings
Special Reporting Requirements
Animal, Human Subjects, Biohazards

We welcome comments on this system