GrantProse INC.
2503 Cravenridge Place
Garner, NC 27529
phone 919.414.5861
www.grantproseinc.com

STUDENTS DISCOVER EXPECTATIONS FOR EVALUATION AND A COMMON LANGUAGE

> *Evaluation is not separate from, or added to, a project, but rather is a part of it from the beginning. Planning, evaluation, and implementation are all parts of a whole, and they work best when they work together* (National Science Foundation [NSF], 2002).

Individuals implementing a grant program are often acquainted with some of the functions and processes of program evaluation; however, the purpose of evaluation and the role it plays in assessing impacts of the grant program can be underestimated or misunderstood. A good understanding of evaluation may require a learning process to develop a shared "language" among members of the project and evaluation teams that integrates the evaluation into the fabric of the project.

The Role and Purpose of Evaluation

Simply put, evaluation is a systematic assessment of a program or project. Information is gathered to answer various questions: Is the project being implemented as designed? What strategies and activities are being implemented? What are the impacts of the project, for instance, in terms of increased knowledge and skills, changes in attitudes, modified behaviors, and changes in affected populations and the larger community? Is the project meeting goals and objectives described in the proposal? If not, why, and have adjustments been made to introduce improvements in program operations?

An evaluation plan is composed of various elements, some that are standard to most evaluation efforts and some that are tailored to each project. Illustrating some of the standard elements, Figure 1 depicts a model that Grant*Prose* has found to be straightforward and useful for organizing an evaluation plan.

Figure 1. A Model for Organizing an Evaluation Plan

Adapted from Brinkerhoff, Brethower, Hluchyj, & Nowakowski (1983)						
Evaluation Questions	Information Collection Plan				Formative Evaluation & Continuous Improvement Plan	Summative Evaluation & Reports Plan
	Methods & Instruments	Type of Data	Data Sources	Data Timelines		

In designing and conducting an evaluation, it is useful to remember the purposes it serves. The Kellogg Foundation describes evaluation as "the consistent, ongoing collection and analysis of information for use in decision making" (W.K. Kellogg Foundation [Kellogg], 1998), and the National Science Foundation gives two reasons for conducting an evaluation: 1) to gather information to help improve the project, and 2) to provide new, unanticipated, insights (NSF, 2002). The latter can have a profound effect on shaping policy (Patton, 2001). An evaluation can have other useful effects, including building capacity of project staff and keeping funding agencies informed.

Methods include both formative (aka process) and summative (aka outcome) evaluations, along with other possible methods such as experimental design (e.g., randomized treatment and control groups, matched controls, case studies). While methods, experimental design, and gathering data are crucial, an evaluation is more than the sum of its parts; it should provide a view "from above" of the entire project—whether from 30,000 or 300 feet—offering the perspective and opportunity for continuous program improvement.

Methods of Evaluation

Formative Evaluation

Formative evaluation, also known as process evaluation, focuses on how a program is being implemented. It describes how the program operates, the services it delivers, and the functions it carries out. Much like monitoring, the formative evaluation addresses whether the program is being implemented with fidelity to its original design and is providing services as intended. When changes are introduced to program operations, formative evaluation documents reasons for the deviation from the original design.

Formative evaluation is typically an ongoing activity, occurring throughout the period of program operations, and usually involves qualitative forms of measurement instrumentation (e.g., interviews, focus groups, observations, attendance at leadership meetings, and review of program records).

Along with monitoring fidelity to the project design, the formative evaluation is a vehicle for organizing and providing feedback on program operations—information useful to introducing refinements and continuous improvements in the program, important to project managers and stakeholders concerned with the project's success.

Summative Evaluation

While formative evaluation addresses a question of how change occurred, summative evaluation, also known as outcome evaluation, measures the nature and/or amount of this change. Summative evaluation measures progress toward the desired goals, objectives and outcomes of the program and identifies the results of a program's effort. It seeks to answer an overarching question, "What difference did the program make?" providing a statement about the net impacts of a program after a specified period of operation.

Summative evaluation typically employs quantitative measurement instrumentation (e.g., surveys, percentages, test scores, period of employment, change in pollution, etc.) to quantify change. Generally speaking, quantitative measures can be added, subtracted, multiplied, and divided; objectives or outcomes are usually addressed in a numerical fashion (e.g., "Among participants in this program, there will be a 10% increase passing their final exam relative to baseline from the previous year").

The summative evaluation is conducted at specific points in time, when data become available, and results of the summative evaluation may be an important determinant of whether the program receives continued funding.

Two Caveats

The Importance of Formative Evaluation

Formative evaluation has multiple purposes and uses, including identifying barriers to implementation, determining whether goals and objectives meet the target population's needs, ascertaining how staff and clients relate to one another, and monitoring stakeholders' experiences with the project, among others (Kellogg, 1998). Formative evaluation may indicate the need for strategies and activities to be "tweaked" in response to local needs and other factors: "Even well-planned projects need to be fine-tuned in the first months of operation, and often information needs to be continually analyzed to make improvements along the way" (Kellogg, 1998, p. 24).

By documenting the program's development and operation, formative evaluation allows an assessment of the reasons for successful or unsuccessful performance, and provides information for potential replication and/or extending operations in additional or different contexts. As such, formative evaluation may be even more important than summative evaluation.

Evaluation vs. Research

Evaluation and research are not the same thing, and it is possible to evaluate a research initiative. The National Institute of General Medical Sciences (n.d.), for instance, indicates *Research is scientific inquiry based on intellectual curiosity, and it produces generalizable knowledge that advances a field. In contrast, evaluation judges the worth or merit of a particular program. It focuses on information for decision-making, taking into account specific program goals and stakeholder interests.* Like research, program evaluation involves measurement, but its emphasis on continuous improvement sets it apart. Evaluation should create a "…balance between proving that programs work and improving how they work" (Kellogg, 1998, p. iii). Program evaluation should provide frequent feedback to the project leaders on how the project appears to be functioning. Evaluation should serve as a management and learning tool for stakeholders, and involve project managers and staff, creating a participatory approach to project improvement (Kellogg, 1998).

The Common Language of Evaluation

Logic modeling often undergirds an evaluation plan. A well-constructed logic model helps identify significant program elements such as activities, outputs, and outcomes (Bureau of Justice Assistance, n.d.) and helps define what to measure during the evaluation (Office of Juvenile Justice and Delinquency Prevention, n.d.). While there are many variations of logic models, most models describe inputs, strategies (and/or activities), outputs, and outcomes.

Inputs are whatever you "put into" the project to implement it, namely, plans and resources (Coffman, 1999). Grant money, representing financial resources, is one input. Others include community resources (programs, organizations, and people), organizational resources (e.g., the building used for the program), and human resources (e.g., leaders, volunteers) (W. K. Kellogg Foundation, 1998/2004).

Strategies and activities are what you do to "make it all happen," plans and actions undertaken to achieve project objectives. Strategies are broad approaches, such as management, capacity building, professional development, or educational programming. Activities are carried out in

support of strategies—they have a beginning and an end, and milestones help measure when an activity is completed. Examples include conducting training workshops, hiring professional speakers, implementing an afterschool program.

Outputs are deliverables and things that result from activities, such as presentations made at conferences, reports disseminated to key stakeholders, curriculum modules developed for academic programs, and individuals completing a training program. Outputs can be counted and may represent a desired impact of the program. Outputs are sometimes confused with outcomes.

Outcomes relate to changes in people created as a result of project activities and outputs. For instance, are 50% more teachers better prepared to teach a relevant subject? Do students demonstrate improved performance on measures of science knowledge? Have there been changes in institutional practices that lead to greater collaboration of institutional members? Outcomes can be considered short term (0-12 months), with changes usually appearing in such realms as knowledge, skills, and attitudes; intermediate (12-60 months), with changes in things like behaviors, practices, and policies; or long term (continuing beyond the grant period), with changes in environmental, social, economic, and civic conditions (University of Wisconsin-Extension, n.d.).

It can sometimes be difficult to distinguish between outputs and outcomes. While it is preferable to maintain a distinction—if there is to be any "logic" in logic modeling—doing so may be less important to the evaluation than determining the impact the outputs and outcomes, however they're characterized, are having on the success of the program. Appendix A provides an example of logic modeling for the Students Discover program.

Relevant Data

There are two main types of data collected during an evaluation:

- Qualitative data, which can be ordinal or nominal. These data are gathered through such methods as interviews, observations, and focus groups.
- Quantitative data, which are discrete or continuous. These data originate from such instruments as enrollment rosters, frequency counts, and test scores.

The distinction can be blurred at times, for instance, with qualitative interviews having items such as multiple choice or yes/no responses that could be quantified. Data sources include whatever helps measure implementation, project activities, and desired outputs and outcomes (e.g, teachers, parents, students, test scores, attendance records, clients, participation logs, archival records, website analytics, and more).

Summary

It is important for those charged with implementing a program and those charged with evaluating the program to have a common understanding of the purposes, processes, and language typically used in the evaluation. In this document, we have discussed the different purposes of formative and summative evaluations, described the process of creating an evaluation plan (as modeled upon Brinkerhoff, et al., 1983), and defined a variety of terms common to evaluation activities (e.g, inputs, strategies, activities, outputs, outcomes, qualitative, quantitative).

Ultimately, an evaluation effort must be grounded in the design and purposes of the program that is being evaluated. Please see Appendix B for source material associated with the Grant*Prose* evaluation of Students Discover.

References

Brinkerhoff, R.O., Brethower, D.M., Hluchyj, T., & Nowakowski, J.R. (1983). *Program evaluation: A practitioner's guide for trainers and educators*. Boston: Kluwer-Nijhoff Publishing

Bureau of Justice Assistance. (n.d.). *Guide to program evaluation.* Retrieved from https://www.bja.gov/evaluation/guide/bja-guide-program-evaluation.pdf

Coffman, J. (1999). *Learning from logic models: An example of a family/school partnership program* (Reaching Results). Retrieved from Harvard Family Research Project website: http://www.hfrp.org/publications-resources/publications-series/reaching-results/learning-from-logic-models-an-example-of-a-family-school-partnership-program

National Institute of General Medical Sciences. *Evaluation basics*. Retrieved from http://www.nigms.nih.gov/Research/Evaluation/evaluation_faqs.htm

National Science Foundation, Division of Research, Evaluation and Communication. (2002). The 2002 user-friendly handbook for project evaluation (Contract REC 99-12175). Retrieved from http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf

Office of Juvenile Justice and Delinquency Prevention. (n.d.). *Performance measures.* Retrieved from http://www.ojjdp.gov/grantees/pm/logic_models_understanding.html

Patton, M. Q. 2001. *American Journal of Evaluation* 22(3). Evaluation, knowledge management, best practices, and high quality lessons learned. 329-336.

University of Wisconsin-Extension. (n.d.). *Outputs vs. outcomes.* Retrieved from http://www.uwex.edu/ces/lmcourse/Module_1_pages/M1_Section2/HTML/m1s2p3a.htm

W. K. Kellogg Foundation. (1998). *W. K. Kellogg Foundation evaluation handbook*. Retrieved from http://www.wkkf.org/knowledge-center/resources/2010/w-k-kellogg-foundation-evaluation-handbook.aspx

W. K. Kellogg Foundation. (1998/2004). *W. K. Kellogg Foundation logic model development guide.* Retrieved from http://www.wkkf.org/knowledge-center/resources/2006/02/wk-kellogg-foundation-logic-model-development-guide.aspx

APPENDIX A

Logic Modeling with the Students Discover Program

Inputs	Strategies	Outputs	Outcomes

• NSF MSP funds
• NSF guidances
• Students Discover leadership
• Facilities and resources
• Partnerships
• MSPnet Hub
	• Management
	• Collaboration
	• Recruitment & retention
	• Professional development
	• Curriculum design
	• Educational instruction
	• Dissemination
	• Research
	  • Basic Sci
	  • Scaling
	• Evaluation
		• Teachers completing Kenan Fellowships
		• Curriculum modules
		• Students completing curriculum modules
		• Research findings
		• Conference presentations
		• Publications
			<u>Short-term (0-12 mos)</u>
			• Changes in teaching practices
			
			<u>Intermediate (12-60 mos)</u>
			• Changes in students' attitudes towards science
			• Changes in students' knowledge of science
			
			<u>Long-term (60+ mos)</u>
			• New scientific knowledge
			• Changes in institutional practices
			• Sustainable and scalable program demonstrated

Of note, the logic model depicted here for the Students Discover program is intended to be dynamic and subject to further development as the Students Discover program unfolds in the coming months.

APPENDIX B

Source Material Associated with Evaluation of the Students Discover Program

NSF MSP Program Solicitation

Requirements of ALL Partnership Proposals - *An external evaluator, evaluation questions, and evaluation design linked to the goals and outcomes of the project; the analysis of data informs the continuous refinement of the project.* (p. 6)

MSP Data Collection, Program Evaluation, Knowledge Management and Dissemination - *The MSP program has funded the development of online data collection modules in an MSP Management Information System (MSP-MIS) to collect common data from funded projects. The program has also awarded an external contract for overall program level evaluation (MSP-PE) that addresses evaluation questions consonant with the role of the MSP program as part of a research and development venture in K-12 STEM education. Thus, the MSP-PE will address evaluation questions not only about the impacts MSP projects might have produced, but also about their contribution to advancing knowledge in STEM education.* (p. 7)

Evaluation Plan - *Formative evaluation should provide evidence of the strengths and weaknesses of the project, informing the Partnership's understanding of what works and what does not in order to inform project progress and success. Summative evaluation should give an objective analysis of qualitative and quantitative data, thus demonstrating the effectiveness of the project on student and teacher outcomes and institutional change among all Core Partners. Although the Evaluation Plan will be developed with input from the Partnership, objective analyses and findings require either an external evaluator or an objective evaluator within a partner institution who is clearly separate and distinct from the partnership participants and their departments/units (e.g., in a department/unit within a university that is not part of the Partnership itself). The qualifications of the evaluator(s) must be provided in the proposal.* (p. 12)

* * * * * * * * *

MSP Proposal Submitted to NSF

The main objectives for the external evaluation of Students Discover are to measure student and teacher outcomes and document institutional change for partners including NCSU, the Museum, and the school districts. Additionally, the evaluation team will document dissemination of the project and assess the potential for future research and use of research results. Formative evaluation will provide an on-going guidance to the project leadership team in terms of progress towards implementation and achieving desired outcomes. The evaluation team will work with the project team to facilitate the design based research effort previously described. Summative evaluation activities will measure changes in student attitudes, knowledge, and course taking behavior, which may be attributed at least partly to project participation. (p. 11).

The following table provides the evaluation questions, instruments/methods, and timeline associated with the four goals described in the NSF MSP proposal.

Evaluation Question	Instrument/Methods	Time
1. Teachers will demonstrate increased use of inquiry practices in science teaching.		
1.1 To what extent do Kenan teachers implement the CS curriculum modules in their classrooms during and after their fellowship year?	• Teacher logs • Website analytics • Research team observations • Student survey	Yrs 2-5
1.2 Does implementation of the CS curriculum result in changes in teaching practices?	• Research and evaluation team observations • Teacher focus groups • Teacher survey	Yrs 2-5
1.3 To what extent do non-KF teachers in participating districts and beyond implement the curriculum modules?	• Teacher logs for those who attend institutes • Website analytics • Research team observations	Yrs 2-5
2. Students will improve STEM attitudes, achievement, and take additional STEM courses.		
2.1 How does KF implementation of curriculum modules affect their students' attitudes towards science and knowledge of both core content and scientific research?	• Student Attitudes towards STEM (FI Middle School STEM Attitudes Survey) pre/post • Project developed instrument • NC 8th grade science End-of-Grade tests	Yrs 2-5
2.2a How does participation in the Summer Bridge and Saturday Academy program affect students' attitudes? 2.2b. Compared to a matched control group, how does participation in the Summer Bridge and Saturday Academy affect students' knowledge and enrollment in subsequent science courses?	• Student Attitudes towards STEM survey (FI Middle School STEM Attitudes Survey) • Project developed instrument of science knowledge • Algebra and Biology End-of-Course assessments • Student science course taking in grades 9-11	Yrs 3-5
3. Partner will demonstrate institutional change in support of students experiencing science.		
3.1a Do all partners show greater awareness of partners' missions, constraints, needs and resources? 3.1b What changes in practices, policies, and procedures are made by partner institutions?	• Surveys and interviews with project management, disciplinary partners, school administrators, Museum staff	Yrs 1-5
3.2 To what extent are KFs able to disseminate the CS projects beyond their classrooms to other teachers, student, parents, and community members?	• Surveys and interviews with project management, disciplinary partners, school administrators, museum staff	Yrs 3-5
4. Project will produce scalable model, attractive curriculum units, and valuable science.		
4.1a To what extent does the project create rigorous challenging and attractive curriculum modules? 4.1b To what extent do the non-KF teachers find the curriculum modules meet their needs and align with NC standards?	• Project developed rubric for rating quality of CS curriculum modules and alignment with NC standards • Teacher survey and interviews	Yrs 2-5

4.2 To what extent do middle school participants produce useful scientific data?	• Survey and interviews with project scientists • Journal publications and conference presentations • Website analytics	Yrs 3-5
4.3 To what extent does the project create and disseminate a sustainable scalable model for joint development of CS projects to promote STEM learning objectives?	• Interviews with co-PIs • Presentations, publications, and spread beyond North Carolina	Yrs 4-5

* * * * * * * * *

NCSU Response to Questions Raised By NSF

In the *Students Discover* grant program, Grant*Prose* evaluation activities differ categorically from the NCSU Friday Institute research effort. The independent evaluation will assess impacts of the project and monitor the fidelity of the grant initiative to its original design. In contrast, the research effort conducted by the NCSU Friday Institute will investigate the process of scaling an educational innovation and contribute to the ongoing iterative design of the Citizen Science (CS) projects. The Grant*Prose* evaluation team functions as the independent, disinterested party, documenting the work processes, outputs, and outcomes produced by the key stakeholders in the project. The evaluation team will also assist in reporting outcomes to NSF and take a lead role in the project's participation in the NSF overall program evaluation.

Grant*Prose* evaluation activities will quantify the outputs and outcomes of the project. As discussed in the *Students Discover* proposal, the Grant*Prose* evaluation team will document:

- Changes in teachers' perceptions, attitudes, and pedagogy over the period of the grant initiative;
- Changes in students' perceptions, attitudes, and science and math achievement over the period of the grant initiative; and
- Changes in institutional practices reported by project leaders including university faculty, museum staff, school administrators, and teachers;

If changes are introduced to the project design, the Grant*Prose* evaluation will document reasons for the change and note what if any impact these changes may have on the original scope and objectives of the grant initiative.