

ARG

Affinity Research Group Model

**Evaluating the Undergraduate Research Experience in Computer Science:
Developing a Framework for Gathering Information about Effectiveness and
Impact**

Evaluating the Undergraduate Research Experience in Computer Science: Developing a Framework for Gathering Information about Effectiveness and Impact

Connie Kubo Della-Piana
Model Institutions for Excellence
connie@utep.edu

University of Texas at El Paso

Andrew Bernat
Department of Computer Science
abernat@cs.utep.edu

Abstract - As undergraduate research experience programs proliferate to include all students, the need to evaluate and assess these types of programs also grows. This paper presents an overview of the evaluation/assessment process and a framework for developing an evaluation plan to assess the effectiveness and impact of the undergraduate research experience. Drawing on the ongoing evaluation of the Affinity Research Group at the University of Texas at El Paso, the framework is designed to (1) focus on issues and questions that form the basis for the design and implementation of the evaluation of undergraduate research experience programs in computer science and (2) provide examples of indicators and tools for assessing program outcomes and process.

Introduction

As the undergraduate research experience is expanded to include all students and funds from private and public institutions increase to support such an effort, the need to evaluate and assess the effectiveness and impact of these programs also grows [1].

No one design, method, or instrument can adequately address all outcomes and processes, therefore this paper provides one framework for developing a plan that guides a systematic examination of programs designed to provide research experiences to undergraduate students. Although this paper only touches briefly on the program goals and the development of an understanding of the program, this step is essential for designing an evaluation plan. The evaluation plan is designed with the "ideal program" in mind and with an understanding of the context in which the program is implemented and its stage of development [2]. Evaluation or the systematic collection of information about an undergraduate research program should provide evidence to (1) render judgment about the overall merit or worth of the research experience; (2) improve the program/project; and (3) generate knowledge about the essential elements of an undergraduate research experience that enhance the effectiveness and impact of such programs/projects.

Before embarking on the arduous task of evaluating an undergraduate research experience program or project in computer science, answers to the following questions need

to be constructed and agreed upon by program developers and evaluators:

- What is the purpose (or are the purposes) of the evaluation?
- What questions about the program need to be answered?
- What information or chain of evidence is needed?
- For whom is the evaluation?
- What led to the decision to evaluate the program/project?
- Who will conduct the evaluation?

After answering the above questions, one proceeds to design the evaluation plan [3]. This paper introduces a framework for developing an evaluation plan for out-of-class activities. Examples from the evaluation of the Affinity Research Group at the University of Texas at El Paso will be used to illustrate the process and implementation of the evaluation plan, indicators and methods to assess the project.

Evaluation

Evaluation is a systematic effort designed and implemented to produce "maximally useful evidence" to either:

1. render judgment about the overall merit or worth of the program;
2. improve the program; or
3. generate knowledge about the issue(s) that the program is designed to address.

The evaluation should provide information that leads to statements about what happened in the original site and what to expect if a certain plan of action is adopted in other sites or continued and adapted in the original site [4].

For example, the overarching evaluation questions that guide the evaluation of the Affinity Research Group are:

1. What impact does participation in the Affinity Research Group project have on student ability to succeed in undergraduate research, to work effectively in groups, and to seriously consider graduate school as an option after receiving their undergraduate degree in computer science?

2. How have activities and the structure of the Affinity Research Group project increased the number of students successfully engaged in research and influenced students who are not typically involved in undergraduate research?

The field of evaluation distinguishes two major types of evaluation: formative evaluation and summative evaluation. Each type or function is conducted for specific and intended uses. Each can contribute to the knowledge base about the issue(s) being addressed by the program. The distinction between formative and summative evaluation is fundamental to program evaluation. On the one hand, formative evaluation is evaluation designed and implemented to support improvement of the program, and is normally commissioned, done by, and delivered to someone who can make improvements. Evaluators work with program developers to come to an understanding about the important questions. The focus of the evaluation is on gathering information about key aspects of the program, such as depicting the strengths and limitations of the program, the nature of its implementation, the level of accomplishment of its intended outcomes or the types of unanticipated consequences that arise.

On the other hand, summative evaluation is done for, or by, any observers or decision makers (by contrast with developers), who need evaluative conclusions for any other reasons besides development [6]. Evaluators may or may not work with program people. Evaluators are more concerned with assessing the worthiness of the program and the impact of the program on participants. A judgment of the value of the program is rendered. Summative evaluation is designed to produce information that can be used to make decisions about the overall success of the program [5].

The distinction between formative evaluation and summative evaluation is illustrated by the following analogy: When the cook tastes the soup, that's formative evaluation; when the guest tastes it, that's summative evaluation [6]. Both are concerned with the processes and outcomes of programs.

In this paper we focus on formative evaluation in order to answer the following questions: Are the activities provided ones that are likely to lead to the attainment of the program/project goals? Is the program/project providing "value-added?" And specifically, do the undergraduate research experiences as designed and implemented in a computer science environment match what is known about best practices in undergraduate research? [7]

Formative evaluation provides a valuable "watch out" service for intended outcomes and unanticipated consequences of participation in the program. The unanticipated consequence of student participation in the Affinity Research Group project is the intensity of recruitment by companies and other faculty for students who have this type of experience. Based on this information, the

project has had to adapt to these external pressures on students by providing students with information on the benefits of continuing on to graduate school and developing a desire and commitment in students to complete the research in which they are involved.

Formative Evaluation and the Undergraduate Research Experience in Computer Science

Evaluation should be guided by questions first and then by the selection of methods, tools and instruments that can help to answer them. Formative evaluation is designed to answer questions such as, "what", "for whom", "by whom", "how", "where", and "why"? Formative evaluation (1) provides data and information to program people on their questions about the program and its implementation and its effect on its participants and (2) presents feedback of the results of data analysis/interpretation to program people for the purpose of on-going improvement of the program.

Table 1 presents examples of some of the types of general evaluation questions that can be asked and specific examples of questions used to evaluate the Affinity Research Group. The table provides a template for generating evaluation questions specific to a program or project [4]. Answers to the questions can help program developers or those who are implementing the program examine their efforts and modify the program's activities or structure to better meet program/project goals.

Design and Implementation

Using Deming's Continuous Improvement Cycle of plan-do-check-act, formative evaluation becomes a systematic and continuous process for asking questions, designing ways to systematically collect information, formulating answers based on the information, deciding how to improve the program and its implementation, revising the program and its implementation, and ... beginning again. The point is that evaluation is not a "one-shot" activity, but an on-going process to inform program developers and those who are implementing program activities. The following description presents a plan of action in the form of questions and answers. Answers to the questions will guide the design and implementation of an evaluation of out-of-class activities, such as an undergraduate research experience that is not course-bound.

The evaluation process travels through the following cycle:

- **Formulation:** What is the purpose of the evaluation and who would be interested in the information?
- **Plan:** What are the overarching questions? What is needed to evaluate the program? How can the relevant information be collected?

- **Do:** How will program developers and evaluators work together to implement the evaluation plan?
- **Check:** How will the information be organized for comparing the "ideal program" and "best practices" to the "program as implemented" and "program practices"?
- **Act:** How will program developers and evaluators work together to utilize the evaluation to improve the program?
- **Start planning again:** How can the evaluation be made more effective in another cycle of planning and implementation?

The outline above provides a strategy for the development of the evaluation effort. Information gathered can be organized in a planning matrix (see Table 2.)

Based on completing the questions a brief statement summarizing the evaluation plan should be created. For example, the following statement was developed for the Affinity Research Group evaluation plan:

The evaluation of the Affinity Research Group Project is a multi-method approach designed to (1) provide information to program developers to improve the conceptualization and implementation of the project and (2) document and examine potential essential organizational structures that define a model for undergraduate/graduate research groups. The evaluation is guided by two issues: (1) what are the important student experiences that enable and encourage student success in computing and (2) what are the essential elements that support student success in computing.

Briefly, the Affinity Research Group evaluation plan includes:

- The administering of a critical incidents technique instrument that is designed to systematically gather and categorize behaviors which have previously been critical to success or failure in a specific performance context or situation.
- The development and administering of a student questionnaire that is designed to systematically gather student self-ratings of their skills and knowledge level in a field of interest and knowledge and skills needed to work in team, as well as, their attitudes toward research and team-building.
- The development and implementation of an organizational ethnography to depict the complexity of an undergraduate experience program/project.

Matching questions with data collection methods

Key to selecting the method or procedure for collecting information is matching the selection of methods to the questions being asked and the information needed within

resource constraints. The advantages and drawbacks of various data collection procedures are explored in [9].

For example, in evaluating the Affinity Research Group, three methods were selected to answer the evaluation questions:

- self-administered questionnaires (self-assessment of skills and knowledge and critical incidents instrument)
- structured interviewing (possible selves, critical events, and exit interview)
- ethnographic observation (intensive at the initial stages of the project)

Table 3 presents an example of how evaluation questions are linked to method and the type of information gathered.

Summary

Evaluation is the systematic collection of information that leads to statements about what has happened and is happening in the original site and what to expect if a certain plan of action is adopted in others sites or continued and adapted in the original site. Evaluation that is useful is guided by the questions needing answers and is sensitive to the evolution of the program, the issues being addressed by the program and the context of program implementation.

In the context of the undergraduate research experience in computer science, particularly out-of-class activities, an understanding of the program is the necessary first step. Once the program is understood by program developers, those implementing the program, and the evaluator a plan to design and implement the evaluation can commence. The paper provides a framework for the development of such a plan based on Deming's plan-do-check-act cycle. The approach to evaluation that is presented in this paper suggests that a collaboration of program people and evaluators be formed and the evaluation undertaken based on the purpose of the evaluation and evaluation questions, rather than on the selection of methods or instrumentation. The goal of the evaluation is to collect information that provides evidence as to the effectiveness and impact of the program and the essential processes and elements that give rise to its effectiveness and impact.

Acknowledgments

This work has been supported by NSF through grants CDA--9522207 and EHR--9550502.

Bibliography

Gates, A., P. Teller, A. Bernat, N. Delgado, and C. Kubo Della-Piana, "Meeting the Challenge of Expanding Participation in the Undergraduate Research Experience",

Proc. of the 1998 Frontiers in Education Conference, 1998, pp. 1133--1138, Phoenix, AZ.
 Cronbach, L. J., S. R. Ambron, S. M. Dornbusch, R. D. Hess, R. C. Hornik, D. C. Phillips, D. F. Walker, and S. S. Weiner, *Toward Reform of Program Evaluation*, San Francisco: Jossey-Bass Publishers, 1980.
 Cronbach, L. J., and K. Shapiro, *Designing Evaluations of Educational and Social Programs*, San Francisco: Jossey-Bass Publishers, 1982.
 Patton, M. Q., *Utilization-Focused Evaluation: The New Century Text*, (Third Edition), Thousand Oaks: Sage Publications, 1997.
 Scriven, M., *Evaluation Thesaurus*, (Fourth Edition), Newbury Park: Sage Publications, 1991.

Chen, H-T, "A Comprehensive Typology for Program Evaluation", *Evaluation Practice*, **17**, No. 2, 1996, pp. 121--130.
 Westat, Inc., *Profiling the Science and Engineering Research Semester: An Approach to Formative Evaluation*, Rockville, MD: Westat, Inc., 1996.
 Besterfield-Sacre, M., and C. Kubo Della-Piana, "An Approach to ABET Accreditation: Formative Evaluation and Continuous Improvement", *Proc. of the Best Assessment Practices in Engineering Education II: A Working Symposium*, October 1998, pp. 1--10.
 Floraline, S., F. Lawrenz, F. Sharp, and J. Frechtling, *User-Friendly Handbook for Project Evaluation: Science, Mathematics, Engineering and Technolgoey Education*, Washington, DC: National Science Foundation, Division of Research, Evaluation and Communication, 1993.

Table 1. Question focus and examples

Focus of the evaluation	Example of general improvement-oriented question	Example evaluation questions for the Affinity Research Group
Program	What are program strengths and weaknesses?	What are the strengths and limitations of the structure of the research experience as implemented in the Affinity Research Group?
Participant progress toward outcomes	To what extent are participants progressing toward the desired outcomes?	To what degree have students increased their understanding of research and its practice? To what degree have students increased their ability to participate in and/or conduct research in computer science?
Treatment--- participant interaction effects	Which types of participants are making good progress and which types aren't doing so well?	How is leaving or continuing with the program a function of student characteristics and program activities, requirements and implementation?
Implementation problems	What kinds of implementation problems have emerged and how are they being addressed?	What variables extraneous to the practice of research exert influence on a student's commitment to research and/or on student continuation in computer science?
Unanticipated consequences	What's happening that wasn't expected?	What happens to a student who is now more confident and competent in computer science because of his/her participation in the program?
Social interaction between staff and clients	How are staff and clients interacting?	What are the essential elements of good student-faculty mentor relationships in a research experience designed from an Affinity Research Group approach?
Perceptions of the program	What are staff and participant perceptions of the program? What do they like? Dislike? Want to change?	What are the characteristics of the Affinity Group Research Project when it is working well? When it is not working well? (from the eyes of the students)
Perceptions of culture and climate	What are perceptions of the program's culture and climate?	What is the nature of the research culture in the Affinity Research Group as a whole? As small research groups?
Influence of external environment	How is the program's external environment affecting internal operations?	How does the need for external funding and deliverables affect the way the Affinity Research Group project is adapted and implemented?
Efficiencies	Where can efficiencies be realized?	Can the training and socialization of new members be allocated to informal situations guided by veteran members?

New ideas/adaptations	What new ideas are emerging that can be tried out and tested?	How does the new automated project management tool affect the completion of student projects?
Allocation of resources	How are funds being used compared to initial expectations?	How is faculty time being used compared to initial expectations?
Value added	Is the program providing value-added?	Are students who participate in the program perceived as better able to meet the demands of group and independent work situations?
Match with best practices	Do the experiences match what is known about best practices?	How do the activities and structures of the Affinity Research Group match with the best practices of undergraduate research experiences?

Table 2. Planning matrix

Evaluation question	Indicators and methods of assessment	Data source	Data collection schedule	Data Analysis Plan	Resources needed & responsible party

Table 3. Evaluation questions, indicators and methods

Evaluation question	Indicators and methods
What are the strengths of the Affinity Research Group project?	Percentage of students planning and pursuing graduate school/degrees - Self-administered questionnaire
How has student understanding of research and its practices increased?	Percent of students reporting that they are confident in their ability to conduct and or design a research project - Self-administered questionnaire and structured interviews
What are the essential elements of a good student-faculty mentor relationship in computer science?	Critical incidents instrument that asks students to describe when program activities worked well and when they did not - Self-administered questionnaire
What are the characteristics of the Affinity Group Research Project when it is working well? When it is not working well? (from the eyes of the students)	Critical incidents instrument that asks students to describe when program activities worked well and when they did not - Self-administered questionnaire
What is the nature of the research culture in the Affinity Research Group as a whole? As small research groups?	Critical incidents instrument that asks students to describe when program activities worked well and when they did not - Self-administered questionnaire
How do the activities and structures of the Affinity Research Group match with the best practices of undergraduate research experiences?	Comparison between program components, comments from students, meetings with program developers and developed lists of essential elements of the undergraduate research experience. Comparison of the Affinity Research Project with the model of graduate and undergraduate research proposed