

## Inside the Black Box—An Implementation Evaluation Case Study

### Abstract

The case study presented in this article is an example of an implementation evaluation. The evaluation investigated significant components of the implementation of a long-term environmental educational program. Direct observation, evaluation-specific survey data, and historical data were used to determine program integrity as identified by adherence to original expectations, dosage, quality of delivery, participant responsiveness, and differentiation from other programs. The evaluation provided key information for replicating and expanding a successful program and exploring areas in which positive changes can be made. The article illustrates how the evaluation methodology that was applied can be useful for other Extension programs.

**Patricia Rector**  
Environmental and  
Resource Management  
Agent  
Rutgers Cooperative  
Extension of Somerset  
and Morris Counties  
Morristown, New  
Jersey  
[rector@njaes.rutgers.edu](mailto:rector@njaes.rutgers.edu)

**Michele Bakacs**  
Environmental and  
Resource Management  
Agent  
Rutgers Cooperative  
Extension of Middlesex  
and Union Counties  
North Brunswick, New  
Jersey  
[bakacs@njaes.rutgers.edu](mailto:bakacs@njaes.rutgers.edu)

**Amy Rowe**  
Environmental and  
Resource Management  
Agent  
Rutgers Cooperative  
Extension of Passaic  
and Essex Counties  
Roseland, New Jersey  
[rowe@njaes.rutgers.edu](mailto:rowe@njaes.rutgers.edu)

**Bruce Barbour**  
Agricultural and  
Resource Management  
Agent  
Rutgers Cooperative  
Extension of Warren  
County  
Belvidere, New Jersey  
[barbour@njaes.rutgers.edu](mailto:barbour@njaes.rutgers.edu)

## Introduction

Succession or expansion of Extension programs may involve the need to replicate existing programs. Regionalization may lead to broadening of existing programs to serve larger geographic areas. New uses of digital technology may require changes to existing programs. These reasons, among others, increase the necessity of understanding how programs work internally in order to replicate success. When there is a need to replicate a complex program, impact reports convey *what* to accomplish but may not provide a detailed blueprint for *how* to repeat the accomplishments that have resulted from the program.

If an Extension program is multilayered, has various coordinators, is conducted across different locations, or involves other complicating factors, having a methodology for conducting an implementation evaluation is helpful. For an evaluation we needed to conduct, we chose to use a methodology described by Duerden and Witt (2012). We present this case study as an example of how the methodology can be used, and we explore the results and benefits of using it. The purpose of our evaluation was to determine how a particular program was operating, whether it was consistent across program locations, and whether it was staying true to the initial goals and objectives developed 10 years earlier. Bush, Mullis, and Mullis (1995) defined "black box evaluations" as those that consider what goes into a program and what comes out of a program without considering what goes on inside a program. An implementation evaluation delves inside the "black box," differing from a typical program

evaluation in that the focus is on evaluating the effectiveness of the method for implementing the program rather than the program outcomes.

The Rutgers Environmental Steward (RES) program (<http://envirostewards.rutgers.edu>) is a certification program that has been offered annually since 2005 and includes a 20-week lecture series, field trips, and a 60-hr internship. The goals of the program are to increase knowledge and public awareness of scientifically based information related to environmental issues and to enable graduates to facilitate positive change in their communities. It is held in multiple counties and has several coordinators. During its first 10 years, the program provided training to 373 participants on various topics, including soil health, climate change, habitat protection, energy conservation, and water resource protection. Participants have completed 166 volunteer internships of 60+ hr each (9,960+ total hr), often under the direction of host agencies (hosts) focused on improving the environment at the local level (e.g., watershed associations). The dollar value of this volunteer time is estimated at \$255,773 (Independent Sector, 2015). As new coordinators of the program, we initiated an implementation evaluation in 2014 to inform efforts to expand the program and reproduce it in new locations (Domitrovich & Greenberg, 2000).

## Methods

The implementation evaluation method described by Duerden and Witt (2012) is based on the concept of program integrity, or how closely the way in which a program is conducted aligns with what was originally intended. An implementation evaluation measures different aspects of implementation (Dane & Schneider, 1998; Duerden & Witt, 2012), including the following elements:

- adherence (how closely the program implementation matches operational expectations),
- dosage (the amount of time given by people in various program roles to achieve the success of the program),
- quality of delivery (the manner in which the program is provided),
- participant responsiveness (how engaged the participants are in the program), and
- program differentiation (how components of the program are unique with respect to other similar programs).

We determined that the method outlined by Duerden and Witt (2012) was best suited for an implementation evaluation of the RES program because these elements could be captured by data sources available to us and each element would provide different insights into how the program was functioning.

We used three data sources to obtain the information that would assist us with the implementation evaluation:

- our direct observation of classroom protocols;
- a 2014 survey that was emailed to all RES program participants, lecturers, and hosts and contained questions specifically designed to help us measure the various elements of the implementation evaluation; and
- a historical data set that had been developed over the length of the program and included data on participants and internships as well as evaluation data for all program locations.

We evaluated program adherence by observing the curriculum being used and comparing it to the original curriculum to determine whether the course work had remained true to the original plan. Beyond comparing curricular implementation to original expectations, we also evaluated program adherence on the basis of whether the program was meeting the original program objectives.

Six questions on the 2014 survey related to dosage. For example, the question "Do you feel you received enough mentoring time from the RES program during your internship project?" was included to provide one measurement of dosage.

Quality of delivery was predominately determined on the basis of 15 questions on the survey; these questions related to the quality of the program, the lecture series, lecturers, interns, guidance from the RES program during internships, and guidance from hosts. A survey question for hosts about whether an intern met or exceeded expectations had the potential for multiple responses because hosts may sponsor different interns in various years. A host was asked to provide an assessment relative to each intern the agency had sponsored. Speaker evaluations conducted in 2013 and 2014 ( $n = 249$ ) to determine ratings for the individual lecturers were part of the historical data set and were used to provide an overall rating for the lecturer component of the program.

Ten questions on the survey related to participant responsiveness. These questions were asked of the participants (e.g., "Do you visit the Rutgers Environmental Steward website?"), the lecturers (e.g., "Do you feel there is a healthy interaction between you and the students?"), and the hosts (e.g., "Did you maintain a relationship with the intern after the project was completed?").

To gather information for evaluating program differentiation, we engaged in personal correspondence and reviews of materials on the Internet. Specifically, we based our determinations on a review of the Rutgers Master Gardener program and web-based research of other similar programs.

If time and resources are available, an implementation evaluation may address other factors that help assess how a program functions (Duerden & Witt, 2012). As a function of our evaluation, we used Fisher's exact test to determine whether there was a significant relationship between program location and number of internships completed. We also applied a Pearson chi-square test of association using only locations having student populations that completed more than three internships to determine whether this factor changed the results. Both analyses were conducted using SPSS version 2.2. A significance level of  $p < .05$  for all statistical tests was considered significant.

## Results

Response rates for the 2014 survey were 35% for RES program participants ( $n = 129$ ), 40% for lecturers ( $n = 15$ ), and 11% for hosts ( $n = 7$ ). Not all participants completed the 2014 survey, but any completed questions from incomplete surveys were accepted for data analysis.

### Program Adherence

The RES program maintained its stated core curriculum at all locations. Across all years and locations, offer of a certificate of completion for finishing the lecture component is standard, and a completed internship is required prior to a participant's becoming a Certified Rutgers Environmental Steward. Also, all classes offered field trips.

In addition to looking at operational expectations, we investigated how successful the RES program has been in achieving the original stated program objectives. Table 1 shows the original objectives, the tool used to determine whether the program is achieving each objective, and the identified method for meeting each objective.

**Table 1.**

Investigation of the Achievement of RES Program Objectives

<b>Objective</b>	<b>Measurement tool</b>	<b>Method of meeting objective</b>
Participants will become knowledgeable in processes of earth, air, water, and biological systems.	Observation	Each RES class offers individual lectures on each of these topic areas.
Participants will be aware of techniques and tools used to monitor and assess health of systems.	Observation	Several classes address monitoring and accessing data sets. Benthic macroinvertebrate monitoring and precipitation monitoring for climate are included in most classes. In 2014, several locations added an online mapping class that uses digital mapping programs for environmental information.
Participants will have an understanding of the research and regulatory infrastructure of state and federal agencies operating in New Jersey that relate to environmental issues.	Observation	No method identified.
Participants will have an introduction to group dynamics and community leadership.	Observation	Two classes on group dynamics were offered at each location every year except 2014.
Participants will recognize the elements of sound science and public	Observation	All classes address the elements of science throughout the program.

policy based on that science.

Participants will have some sense of the limits of the current understanding of the environment.

Graduates will use their knowledge to create positive change in their communities.

Observation

A lecture on "the limits of science" has been offered since the beginning of the program. In 2014, some evening classes were unable to offer this lecture.

Rutgers Environmental Stewards website

As of 2014, at least 9,960 volunteer hr have been provided to New Jersey communities through internships for environmental improvement projects.

2014 Survey (n = 107)

39% are or have been members of an environmental commission, a planning board, or another municipal, county, or state board or commission.

2014 Survey (n = 42)

Of respondents who were or had been on a board or commission, 64% were not members prior to participating in the RES program.

2014 Survey (n = 15)

Of those who were a member of a board or commission prior to participating in the RES program, 93% felt that the training helped them serve in a more productive way. Of those who became members of commissions after participating in the program, 96% felt that the training helped equip them for the associated responsibilities.

## Dosage

To measure dosage, we relied on past participants, lecturers, and hosts to assess whether sufficient time was allotted for them to successfully complete their tasks. Table 2 summarizes the survey questions and responses related to dosage.

**Table 2.**

Dosage Responses from the 2014 Survey

Survey respondent	Question	% responding		
		"yes"	"no"	"I do not know"

Participant ( <i>n</i> = 38)	Do you feel that you received enough mentoring time from your host agency during your internship project?	71	18	11
Participant ( <i>n</i> = 45)	Do you feel you received enough mentoring time from RES program during your internship project?	73	27	
Lecturer ( <i>n</i> = 11)	Do you feel you received sufficient interaction time with the RES program to successfully prepare your lecture?	91	9	
Host ( <i>n</i> = 14)	Do you feel that you needed more resources or interactions with Rutgers during the internship(s)?	29	71	
Participant ( <i>n</i> = 114)	Did you complete an internship? [related to number of participants who received the full dose of the program (lecture + internship)]	47	53	

## Quality of Delivery

Results of the 2014 survey indicated that the RES program exceeded (61%) or met (37%) the participants' expectations (*n* = 112). A majority of the participants (98%) would recommend the RES program to a friend or colleague (*n* = 112). On program evaluations conducted in 2013 and 2014, participants (*n* = 249) used a rating scale of 1 to 5, with 5 representing *excellent*, to rate the program lecturers; participants' ratings averaged 4.5 for clearly explaining subject matter and 4.4 for overall presentation content.

The majority of RES program hosts (64%) stated that their interns exceeded their expectations, and 29% stated that their interns met their expectations (*n* = 14). All the host respondents to the 2014 survey said they would host an intern again (*n* = 7). See Table 3 for additional survey questions relating to quality of delivery.

**Table 3.**

Additional Questions for Obtaining Information on Quality of Delivery

Question	%	%	%
	responding "very"	responding "somewhat"	responding "not at all"
When reflecting on the lecture series, how	95	5	

successful was the lecture series as a means to increase your environmental knowledge? (*n* = 112)

Please rate how helpful the quality of the guidance from your host agency was during your internship project. ( <i>n</i> = 37)	64	23	13
--	----	----	----

Please rate how helpful the quality of the guidance from the Rutgers Environmental Steward Program was during your internship project. ( <i>n</i> = 58)	46	46	9
---	----	----	---

---

## Participant Responsiveness

Lecturers' responses to the 2014 survey indicated that they perceived a healthy interaction between themselves and their audiences (91%; *n* = 15). This finding is important because vigorous interaction between a program's lecturers and its participants is a useful indicator of participant responsiveness.

Another important measure of participant responsiveness is whether a participant completed an internship and became certified. For survey respondents who had completed an internship (47%), a majority (61%) worked with a host agency. Additionally, interns had maintained relationships with their host agencies; 86% of host agencies indicated that they had had further interactions with their interns (*n* = 7).

Of the program participant survey respondents who did not complete an internship, 61% said they were not sure what to do or could not get an internship project off the ground, and the remainder cited personal reasons, such as family obligations or lack of time. Thirty-one percent responded that they had started an internship but did not complete it.

## Program Differentiation

Many Extension programs across the country are similar to the RES program in that they provide education in return for requested or required volunteer hours (e.g., the Georgia Master Naturalist and Mississippi Master Naturalist programs). Within New Jersey, the Rutgers Master Gardener program is a statewide program, offered on a county basis, with weekly expert lectures and required internships. Table 4 summarizes how the RES program is similar to and different from other Extension volunteer programs.

**Table 4.**

Similar Programs and Ways in Which the RES Program Is Similar or Dissimilar to Those Programs

Program	Program component	Specifics	RES program
Rutgers Master Gardener program	Lecture	Subject matter experts lecturing each week on a different topic	Similar
	Volunteer service	Volunteer hours split across programming (helpline, demonstration projects, standing commitments at local gardens)	Dissimilar; 60 hr on one focused, usually self-

			directed project
Georgia Master Naturalist program	Lecture	Combination of lecture and outside hands-on learning	Similar; RES not focused on outside learning
	Volunteer service	Encouraged to share their knowledge with their communities by volunteering at local schools or nature centers	Dissimilar; RES has a requirement to complete volunteer service
Mississippi Master Naturalist program	Lecture	Subject matter experts with 40 hr training	Similar
	Volunteer service	Required to complete 40 hr service within 12 months of completion of lecture; service through a variety of activities, such as educational activities, projects, and demonstrations	Dissimilar; RES internship is required to be focused on only one project

## Location, Location, Location

Across program locations, there were differences in average number of internships completed. Information about program locations and average numbers of internships completed is shown in Table 5.

**Table 5.**

Number of Years Program Was Conducted and Number of Internships Completed per Year by Location

Location	Number of years program was conducted (2005–2014)	Average number of internships completed per year
Gloucester County	1	6
Burlington County	5	5.2
Essex County	5	4.8
Somerset County	9	4.8
Atlantic County	6	2.5
Warren County	2	2
Middlesex County	1	1
Passaic County	1	1
Cape May County	1	0

Statistical analysis indicated that the variation in the numbers of completed internships across program locations



is significant (Fisher's exact = 22.4,  $p = .003$ ). A separate analysis that excluded locations with three or fewer completed internships also indicated that the variation in the numbers of internships across program locations is significant ( $\chi^2 = 11.9$ ,  $df = 5$ ,  $p = .035$ ).

## Discussion

Our evaluation indicated that the program, overall, had adhered to its original intent and objectives over 10 years and across various locations. The process of implementing the RES program relies heavily on the lecturers who contribute to the lecture series. The lecturers, as a group, were rated as very good across all years and all locations. Also, the lecture series, overall, was functioning well and had maintained program integrity.

Internships were identified as an area in which program integrity, although maintained in theory, was floundering in process. The evaluation enabled us to go further than simply identifying a problem. Starting the evaluation by examining program adherence showed us that the original requirements remained unchanged; therefore, the problems were in implementation of the internship program.

Through evaluation of dosage, we found that a substantial number of participants were not receiving the full dose of the program (lecture plus internship) and that some participants felt they may not have received the time they needed from others to help them accomplish their internships. Lack of sufficient personnel can affect program implementation and is often a factor with nonimplementers (Kramer, Lauman, & Brunson, 2000).

Data related to quality of delivery told us, first, that the program was considered high quality by the majority of those involved with it. Also, according to hosts, the majority of program interns who worked on projects with them did good jobs.

Program participation responses indicated that hosts and interns maintained relationships after internship projects were complete. Positive relationships between hosts and interns and high internship completion rates among interns who worked with host agencies seem to indicate that linking interns and hosts may help increase the ability of program participants to take part fully in the program.

Considering program differentiation helped us note differences between our internship program and other volunteer programs. We are in the process of looking for partial solutions to issues we identified by exploring this element. One idea is development of an alternative track for internships. We could offer a resource-rich "canned" opportunity while still maintaining the 60-hr project-focused internship, with additional hand-holding for those who may not be comfortable with a self-directed project.

We can seek to understand the differences among program locations with respect to internships and look for ways to assist counties in which internship completion rates are lower. Additional faculty now assist with the RES program. We are investigating the potential of having a program coordinator help all counties manage the important follow-up that leads to completed internships, perhaps thereby increasing internship completion rates. We can begin to look at solutions related to information we gleaned about program participation, dosage, program differentiation, and location. This potentially broad base for solutions provides more room to effect change within the program.

## Conclusion

We chose to evaluate the RES program implementation process—"how" the program was being implemented—so that as new coordinators we could better replicate its success and work to enhance areas needing improvement.

We based our evaluation on the "program integrity" evaluation method detailed by Duerden and Witt (2012). To conduct our evaluation, we developed strategies to provide measurements for all five elements associated with that evaluation method and obtained data from as many sources as possible (Dane & Schneider, 1998). We looked at the program through the lenses of direct observation, a survey developed specifically to help answer questions related to the evaluation, and historical data. Although the evaluation required time and effort, we feel that the time and effort were well spent and that the evaluation method we used provided an objective means for assessing the internal functioning of the program. For this case study, the evaluation based on the work of Duerden and Witt (2012) provided an excellent means for evaluating the mechanics (implementation system) of a complex program. This method can be modified and used by others addressing programs that are going through changes, need succession plans, span multiple locations, or simply could benefit from a thorough examination.

### Acknowledgments

We would like to acknowledge Kevin Sullivan, assistant director of statistical analysis, Office of Research Analytics, Rutgers, New Jersey Agricultural Experiment Station, for conducting the statistical analysis.

### References

- Bush, C., Mullis, R., & Mullis, A. (1995). Evaluation: An afterthought or an integral part of program development. *Journal of Extension* [online] 33(2) Article 2FEA4. Available at: <http://www.joe.org/joe/1995april/a4.php>
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23–45.
- Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation*, 11(2), 193–221.
- Duerden, M. D., & Witt, P. A. (2012). Assessing program implementation: What it is, why it's important, and how to do it. *Journal of Extension* [online], 50(1) Article 1FEA4. Available at: <http://www.joe.org/joe/2012february/a4.php>
- Independent Sector (2015). *National value of volunteer time*. Retrieved from <http://independentsector.org/uploads/resources/Value-of-Volunteer-Time-by-State-2001-2014.pdf>
- Kramer, L., Laumann, G., & Brunson, L. (2000). Implementation and diffusion of the rainbows program in rural communities: Implications for school-based prevention programming. *Journal of Educational and Psychological Consultation*, 11(1), 37–64.

*Copyright* © by Extension Journal, Inc. ISSN 1077-5315. Articles appearing in the Journal become the property of the Journal. Single copies of articles may be reproduced in electronic or print form for use in educational or training activities. Inclusion of articles in other publications, electronic sources, or systematic large-scale distribution may be done only with prior electronic or written permission of the [Journal Editorial Office, joe-ed@joe.org](mailto:joe-ed@joe.org).

If you have difficulties viewing or printing this page, please contact [JOE Technical Support](#)