

## A National Perspective on the Current Evaluation Activities in Extension

**Alexa J. Lamm**

Director

National Public Policy Evaluation Center

Gainesville, Florida

[alamm@ufl.edu](mailto:alamm@ufl.edu)

**Glenn D. Israel**

Professor

University of Florida

Gainesville, Florida

[gdisrael@ufl.edu](mailto:gdisrael@ufl.edu)

**David Diehl**

Assistant Professor

University of Florida

Gainesville, Florida

[dcdiehl@ufl.edu](mailto:dcdiehl@ufl.edu)

---

**Abstract:** *In order to enhance Extension evaluation efforts it is important to understand current practices. The study reported here researched the evaluation behaviors of county-based Extension professionals. Extension professionals from eight states (n = 1,173) responded to a survey regarding their evaluation data collection, analysis, and reporting methods while evaluating their "best" Extension program. The majority used post-tests at the conclusion of their educational activities and reported the actual numbers of respondents. Few respondents collected data representing behavior change. To improve, Extension professionals need to gain evaluation skills and work together to develop strong evaluations that measure change over time.*

---

### Introduction

The national Cooperative Extension System (Extension) offers some unique challenges when addressing evaluation concerns (Braverman & Engle, 2009). Franz and Townson (2008) stated "complex funding, staffing, and accountability structures combined with widely varying programs and delivery methods make program development and evaluation challenging for Extension" (p. 5). While grants from public agencies and private organizations assist in the development and delivery of unique Extension programs, the majority of funding for Extension as a whole comes from local, state, and federal funds (United States Department of Agriculture National Institute of Food and Agriculture, 2011). The need for Extension to exhibit public value of its programs is increasing as county and state budgets are cut (Chapman-Novakofski et al., 1997; Radhakrishna & Relado, 2009). The ability to provide credible information to the public when making funding choices depends primarily on the evaluation activities of Extension professionals.

Evaluation is one way to measure, examine, and report perceived value or programmatic impact. Evaluation has been defined multiple times as systematically determining something's merit, worth, value, quality, or significance (Davidson, 2005; Fournier, 2005; Schwandt, 2002; Stufflebeam, 2001). Patton (2008) described evaluation as "assessing what was intended (goals and objectives), what happened that was unintended, what was actually implemented, and what outcomes and results were achieved" (p. 5). Rossi, Lipsey, and Freeman (2004) defined program evaluation as "the use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social action in ways that improve social conditions" (p. 431).

Evaluation has always been a part of Extension programming. However, most Extension professionals view it as a "necessary evil" in the

reporting process rather than as an opportunity to identify and document accomplishments (Agnew & Foster, 1991). Unfortunately, funding for Extension programs has become limited, and accountability has become essential for organizational survival (Franz & Townson, 2008). Without proving impact, government programs can become obsolete in the eyes of its citizens, and Extension is no exception (Anderson & Feder, 2007).

In 1977 the Food and Agriculture Act mandated that the Secretary of Agriculture examine the social and economic impacts of Extension programs. Released in 1980, the report found the accountability work of Extension "short on impacts" (Warner & Christenson, 1984, p. 17). Andrews (1983) stated "In an era of accountability, Extension must be able to defend who and how people are being served. It also needs to document that programs are achieving positive results" (p. 8). Shortly after these findings were released, the General Accounting Office also criticized Extension publicly for not having a defined focus (United States General Accounting Office, 1981). This report specifically mentioned the need for improved evaluation and accountability.

In 1993, Congress passed the Government Performance Results Act (GPRA) as a result of a lack of confidence the American public exhibited towards federal funding choices. Goal setting, performance appraisal, and public accountability were brought to the forefront in order to improve the public's confidence (United States Department of Agriculture, 1993). The GPRA promoted a focus on results, service quality, and customer satisfaction. The goal was to provide leaders within large government organizations with information about program quality and effectiveness. The Agricultural Research, Extension, and Education Reform Act (AREERA) of 1998 extended GPRA by requiring approved Plans of Work for Extension and research programs, along with annual reports showing medium and long-term impacts, in order to receive federal funding. Thus, the level of accountability within state Extension systems had to be raised to continue receiving federal funding.

With the passing of the GPRA and AREERA, evaluation capacity building within state Extension organizations was added to the professional development agenda (Arnold, 2002; Franz & Townson, 2008; Rennekamp & Arnold, 2009). However, evaluation efforts have been minimally enhanced, and, on a federal level, Extension continues to generate and report basic information on contacts made and reactions to programs, rather than on behavior changes (medium-term) and social, economic, and environmental (SEE) condition (long-term) changes (Franz & Townson, 2008). On the other hand, many states either encourage or require Extension professionals to use logic modeling when developing plans of work. The three outcomes that must be detailed within a fully developed logic model are short-, medium-, and long-term outcomes (Taylor-Powell, 2005). Therefore, Extension professionals should have their desired outcomes identified each year that can be used drive their evaluate plans and methods.

Without enhanced evaluations, state and federal Extension organizations will continue to be inadequate at reporting programmatic successes at the medium- and long-term level. Extension programs will run the risk of having a lower perceived public value without data proving the medium and long-term changes they create (Anderson & Feder, 2007). In order to meet accountability expectations, Extension professionals must collect, analyze, and use high-quality evaluation data. When the skills to do so are lacking, then appropriate professional development should be provided. Thus, a full understanding of how Extension professionals are currently evaluating their programs, including the identification of weaknesses and strengths, is necessary to know where to begin when attempting to enhance evaluation practices.

## **Purpose and Research Objectives**

The purpose of the study reported here was to describe the evaluation behaviors that county-based Extension professionals are currently practicing. The research objectives were:

1. To determine how Extension professionals are collecting evaluation data that can be used to measure short-, medium-, and long-term outcomes.
2. To determine the data analysis and reporting methods Extension professionals use when evaluating their programs to measure short-, medium-, and long-term outcomes.

## **Methods**

### **Respondents**

The targeted respondents for the study included the 1,795 county-based Extension professionals employed by the University of Arizona, University of Florida, University of Maine, University of Maryland, Montana State University, University of Nebraska, North Carolina State University, and University of Wisconsin in 2010. States were selected to represent variation in Extension's organizational size, its structure (e.g., multi-county versus single-county agents), geography (i.e., all of the USDA regions), and hiring process (whether or not there was a tenure process for Extension professionals). Of the 1,223 responses, 1,173 were considered complete (a response rate of 65.3%) and subsequently used in the analysis.

To address non-response error, differences in gender were examined between respondents and non-respondents. Gender was the demographic variable the researchers were able to obtain for the entire population and allowed for the comparison. A Chi-square test was

run to determine if any significant differences existed between respondents and non-respondents. Differences in gender were non-significant ( $\chi^2 = 1.72$ ,  $p = .19$ ) based on a  $p$  value of  $<.05$  established *a priori*.

### **Instrumentation**

Due to a lack of research on the topic, an instrument measuring Extension professionals' engagement in evaluation behaviors was not available. Therefore, the researchers created an instrument. The instrument requested that respondents answer a series of questions related to how they evaluated their "best" or "most important" Extension program during the previous year. In the study an educational Extension "program" was defined as multiple activities (workshops, field trips, sessions, etc.) that are all planned to work towards common objectives that have been clearly defined (Patton, 2008). Examples include the Master Gardener program, a financial management program, a teen leadership program, or a small acreage management program. An Extension "activity" was defined as a one-time educational experience.

Respondents answered by marking whether or not they had engaged in 28 specific evaluation-focused data collection, data analysis, and reporting methods while evaluating their "best" or "most important" Extension program the previous year. The evaluation methods included were based on a list of essential competencies for program evaluators provided by Ghere, King, Stevahn, and Minnema (2006). The instrument also included a series of questions regarding the participant's personal and professional characteristics. The instrument was reviewed by a panel of experts specializing in survey design, evaluation, and familiarity with Extension organizations for internal validity. A pilot test was conducted with Extension professionals employed by Colorado State University Extension for reliability purposes. Follow-up interviews with a sampling of the pilot respondents were conducted to ensure respondents were able to answer the questions accurately.

### **Procedure**

Prior to any data collection, a protocol, including the informed consent and questionnaire, was submitted to the University of Florida's Institutional Review Board (IRB) for each phase of the study. The informed consent form described the study, detailed how much time it would take to complete the instrument, acknowledged known risks and benefits, and specified the study as voluntary. The protocol was approved (Protocol #2010-U-0531).

The survey for the study was distributed using online software. Extension professionals' use of email as a communication tool enabled the use of an online survey (Dillman, Smyth, & Christian, 2009). The researchers implemented the survey by contacting the targeted respondents via email using Dillman et al.'s (2009) Tailored Design Method. Implementation included an initial request to complete the survey and weekly reminders for 4 weeks. A list of Extension professionals' names, email addresses, and genders was generated by each state system separately. The researcher combined the state lists to create a database of 1,795 Extension professionals. Each participant was assigned a number to ensure confidentiality. All correspondence related to the actual survey were sent by the researcher with approval from each state's Extension administration.

### **Data Analysis**

Quantitative descriptive research methods were used to achieve the research objectives. Data analysis was completed using SPSS18 statistical software package for Windows.

## **Results**

There were 751 female (64.0%) and 422 male (36.0%) respondents. Most of the respondents were Caucasian/White (87.6%,  $n = 1027$ ), with African Americans representing 4.1% ( $n = 48$ ). Hispanic, Native American, and Other categories were represented minimally. The majority of respondents had obtained a Master's degree (70.1%,  $n = 822$ ), while 19.0% ( $n = 223$ ) had obtained a Bachelor's degree. All Extension programmatic areas were represented, with 27.1% ( $n = 318$ ) of respondents focusing on Family and Consumer Sciences/Nutrition, 23.4% ( $n = 275$ ) on 4-H Youth Development, 24.6% ( $n = 289$ ) on Agriculture, and 11.2% ( $n = 131$ ) on Horticulture. The remaining respondents (13.7%) focused their Extension programming in Natural Resources, Community Development, and Sea Grant areas. Almost half of the respondents (43.1%,  $n = 505$ ) were in tenure-tracked positions, and 62.6% ( $n = 316$ ) of the 43.1% who were in tenure-tracked positions had achieved tenure.

### **Data Collection Methods**

When the Extension professionals' evaluation behaviors were examined, it was discovered that 13.6% ( $n = 163$ ) did not engage in the practice of evaluation. Of the respondents using data collection methods, it was found the majority of respondents kept program participation records (82.4%), tracked their respondents' gender (71.7%), used post-tests to evaluate specific activities (70.8%), tracked their respondents' race/ethnicity (68.6%), and conducted interviews to evaluate their activities (64.8%) (Table 1). While it is a positive result that data were being collected in some capacity, less than half (49.3%) of the Extension professionals reported collecting data on the

behavior changes resulting from their programs. In addition, only 37.4% were using interviews to collect data on how their programs were making SEE condition changes.

**Table 1.**  
Respondents' Data Collection Methods (N = 1173)

	n	%
Keep program participation records	96682.4	
Track respondents' gender	84171.7	
Post test to evaluate activities	83070.8	
Track respondents' race/ethnicity	80568.6	
Interviews to evaluate activities	76064.8	
Interviews to evaluate entire program	69959.6	
Post test to evaluate entire program	68258.1	
Interview to evaluate behavior change	57849.3	
Collect artifacts such as press releases, journal publications, and/or Extension fact sheets resulting from the program	56448.1	
Conduct both pre tests and post tests to evaluate single activities	52744.9	
Participant written accounts such as journals and/or log books	51944.2	
Interviews to evaluate SEE changes	43937.4	
Pre/post test to evaluate entire program rather than individual activity	40634.6	
Test to evaluate behavior change	38332.7	
Test to evaluate SEE changes	23620.1	
Comparison group used as a control	60	5.1

### Data Analysis and Reporting Methods

Regarding reviewing how the respondents were analyzing and reporting their results, most were simply reporting the actual number of participants who attended their activities/program (82.4%) (Table 2). Moving a step beyond this basic measurement, 54.7% were reporting means and percentages of responses to specific evaluation items. In addition, summaries of written accounts and artifacts being collected as a result of an Extension program were being reported by approximately half of the respondents. Very few respondents reported standard deviations (11.3%), a statistical technique used to show how much variation exists in participant responses. Even fewer respondents compared groups to see if the respondents in their programs were learning more than someone who did not attend or participate in their program (8.5%, Table 2). In addition, very few respondents reported using any type of inferential statistics (2.0%) that can be used to examine statistical differences over time, to identify if covariates are impacting results, or for prediction purposes to further explain results.

**Table 2.**  
Respondents' Data Analysis and Reporting Methods (N = 1173)

	n	%
Report actual numbers	96682.4	
Summary of written accounts	65756.0	
Report means or percentages	64254.7	
Summary of artifacts collected	63454.0	
Summary of interview results	53145.3	
Examine change over time	32828.0	
Comparing content of interviews for similarities and differences	32828.0	
Member checking interview results	30425.9	
Other form of data analysis and/or reporting	19016.2	
Report standard deviations	13311.3	
Compare groups	100	8.5
Advanced inferential statistics	23	2.0

### Conclusions and Discussion

The results show Extension professionals were evaluating their programs. However, the majority were only using post-tests given at the conclusion of their educational activities to assess their level of success. Because data showing behavior and SEE condition changes must be collected over time, this type of evaluation makes reporting medium- and long-term impacts difficult. These findings confirm Franz and

Townson's (2008) claim that Extension continues to generate and report basic information on contacts made and reactions to programs, rather than on behavior changes (medium-term) and SEE condition (long-term) changes. Creating evaluation plans that measure medium- and long-term impacts is difficult and time consuming but necessary to show the type of impact needed to enhance the perceived public value of Extension programs beyond that required by federal reports (AREERA, 1998).

It needs to be acknowledged that some Extension professionals may lack the expertise to perform evaluation behaviors that measure long-term change or conduct advanced inferential statistics. An implication for Extension administrators is that county-based professionals need to be engaged in professional development opportunities that focus on building the skill sets needed to measure medium- and long-term change. Professional development efforts of this type include creating detailed logic models that connect programming with evaluation to measures specific outcomes, understanding how to develop instruments that measure behavior and SEE condition changes, and learning how to conduct data analysis that can show change over time (Arnold, 2006).

In addition, evaluations that measure behavior and SEE condition changes may be more attainable when working with a group rather than as a single Extension professional due to the time-intensive nature of this type of evaluation (Lamm, Harder, Israel, & Diehl, 2011).

Another implication for Extension is that Extension professionals could consider devising program plans as a team so that they address the same specific measureable outcomes. By working together, the task of creating detailed plans, establishing instruments to measure behavior and SEE condition changes, and conducting data analysis for reporting purposes can be shared across the group, reducing the pressure and time commitment felt by a single individual. As Extension professionals gain evaluation skills and abilities, and work together to develop strong evaluations that measure participant change over time, Extension will be able to further document the true public value of its programs.

## References

Agnew, D. M., & Foster, R. (1991). National trends in programming, preparation and staffing of county level Cooperative Extension service offices as identified by state Extension directors. *Journal of Agricultural Education*, 32, 47-53.

Agricultural Research, Extension, and Education Reform Act. (1998). Public Law 105-185. Washington, DC.

Anderson, J. R., & Feder, G. (2007). Agricultural extension. *Handbook of Agricultural Economics*, 3, 2344-2367.

Andrews, M. (1983). Evaluation: An essential process. *Journal of Extension* [On-line], 21(5). Available at:

<http://www.joe.org/joe/1983september/83-5-a1.pdf>

Arnold, M. E. (2002). Be "logical" about program evaluation: Begin with learning assessment. *Journal of Extension* [On-line], 40(3) Article 3FEA4. Available at: <http://www.joe.org/joe/2002june/a4.php>

Arnold, M. E. (2006). Developing evaluation capacity in extension 4-H field faculty. *American Journal of Evaluation*, 27(2), 257-269.

Braverman, M. T., & Engle, M. (2009). Theory and rigor in Extension program evaluation planning. *Journal of Extension* [On-line], 47(3) Article 3FEA1. Available at: <http://www.joe.org/joe/2009june/a1.php>

Chapman-Novakofski, K., et al. (1997). Evaluating evaluation: What we've learned. *Journal of Extension* [On-line], 35(1) Article IRIB2. Available at: <http://www.joe.org/joe/1997february/rb2.php>

Davidson, J. E. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage Inc.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method* (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Fournier, D. M. (2005). Evaluation. In S. Mathison (Eds.). *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Inc., 139-140.

Franz, N. K., & Townson, L. (2008). The nature of complex organizations: The case of Cooperative Extension. In M.T. Braverman, M. Engle, M.E. Arnold, & R.A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension*. *New Directions for Evaluation*, 120, 5-14.

Ghere, G., King, J. A., Stevahn, L., & Minnema, J. (2006). A professional development unit for reflecting on program evaluator competencies. *American Journal of Evaluation*, 27, 108-123.

Lamm, A. J., Harder, A., Israel, G. D., & Diehl, D. (2011). *Team-based evaluation of Extension programs*. EDIS Publication #WC118. Gainesville, FL: University of Florida. Retrieved from: <http://edis.ifas.ufl.edu/wc118>

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Radhakrishna, R. B., & Relado, R. Z. (2009). A framework to link evaluation questions to program outcomes. *Journal of Extension* [On-line], 47(3) Article 3TOT2. Available at: <http://www.joe.org/joe/2009june/tt2.php>

Rennekamp, R. A., & Arnold, M. E. (2009). What progress, program evaluation? Reflections on a quarter-century of Extension evaluation

practice. *Journal of Extension* [On-line], 47(3) Article 3COM1. Retrieved from <http://www.joe.org/joe/2009june/comm1.php>

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage Publications.

Schwandt, T. A. (2002). *Evaluation practice reconsidered*. New York, NY: Peter Lang.

Stufflebeam, D. L. (2001). Evaluation models. *New Directions for Evaluation*, 89, 7-98.

United States Department of Agriculture (1993). *The Government Performance and Results Act of 1993*. Washington, DC. Retrieved from: [http://www.whitehouse.gov/omb/mgmt-gpra\\_gplaw2m/](http://www.whitehouse.gov/omb/mgmt-gpra_gplaw2m/)

United States Department of Agriculture National Institute of Food and Agriculture. (2011). *About us*. Washington, D.C. Retrieved from: <http://www.csrees.usda.gov/Extension/>

United States General Accounting Office. (1981). *Cooperative Extension Service's mission and federal role need Congressional clarification*. CED-81-119. Washington, DC.

Warner, P. D., & Christenson, J. A. (1984). *The Cooperative Extension Service: A national assessment*. Boulder, CO: Westview Press.

---

*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](#)

© Copyright by Extension  
Journal, Inc. ISSN 1077-5315.  
[Copyright Policy](#)