

# Evaluation Situations, Stakeholders & Strategies<sup>1</sup>

Glenn Israel, David Diehl, and Sebastian Galindo-Gonzalez<sup>2</sup>

New UF/IFAS Extension faculty often ask, “What should I do to evaluate my Extension programs?”, “How much time and money should I spend on evaluation?” and “How much rigor should my evaluation have?” The answer to these questions involves consideration of several factors and, consequently, there is not a one-size-fits-all evaluation strategy. In this fact sheet, we provide guidelines for selecting evaluation strategies that are appropriate for different situations. Our goal is to help extension faculty to tailor their evaluation activities to balance their available time and resources with the situation, as well as their individual and stakeholders’ needs.

## Guideline #1: Walk the talk—use university research to design and conduct “best practices” evaluations

Extension professionals often take pride in bringing university research to help address people’s problems. But we must “walk the talk” and practice what we preach. Using research-based evaluation methods can help save time and money while achieving more rigorous and credible evaluations. Extension professionals are under increasing pressure to deliver programs that lead to significant behavioral changes that benefit society. Funders and stakeholders are asking for evidence that extension programs are making a difference in people’s lives and in their communities. As a result, extension faculty must be able to provide credible

information about the ways in which their programs are leading to behavior changes and community benefits. To provide the strongest evidence possible, evaluations need to be as rigorous and credible as possible by:

- Employing reliable and valid measures. Having good data on knowledge gain might be more useful than poor data on behavior change if the former is viewed to be trustworthy and credible while stakeholders discount the latter. We believe that having high quality measures is the heart and soul of any meaningful evaluation.
- Using a comparison group of nonparticipants. Use of a comparison group allows firmer conclusions about the impact because the group provides a benchmark for an extension program.

In most cases, measuring behavior change requires that data be collected some time after conducting the extension program, in addition to the collection of baseline (pre-program) measures. In some cases, it may be feasible and appropriate to collect follow-up data at the end of a multi-session program that occurred over a two or three month period. We also recognize that some situations call for simple evaluations, such as a customer satisfaction survey or a post-only (end of the meeting) survey. But even simple evaluations can be done poorly or well, depending on whether the appropriate research base is used in formulating the instruments and data collection process.

1. This document is WC090, one of a series of the Agricultural Education and Communication Department, UF/IFAS Extension. Original publication date November 2009. Reviewed April 2019. Visit the EDIS website at <https://edis.ifas.ufl.edu> for the currently supported version of this publication.

2. Glenn Israel, professor, Department of Agricultural Education and Communication; David Diehl, assistant professor, Department of Family, Youth and Community Sciences; and Sebastian Galindo-Gonzalez, assistant in, Department of Agricultural Education and Communication; UF/IFAS Extension, Gainesville, FL 32611. The authors thank Amy Harder, Paul Monaghan, and Pete Vergot for comments on an earlier version.

## Guideline #2: Match the rigor of the evaluation to the importance and intensity of the programmatic effort

Table 1 shows a series of situations ranging from low programmatic importance and intensity to high importance and intensity. Selecting a strategy that matches the situation is the first step in tailoring an evaluation.

Programmatic activities that are relatively low intensity, such as reactive responses to client phone calls, e-mails, and office consultations and proactive one-time presentations and train-the-trainer programs warrant lower levels of time and effort to evaluate these. While train-the-trainer events are often part of a larger extension program, evaluating only the training is similar to conducting a student course evaluation or a customer satisfaction survey. For these types of evaluations, generic surveys developed for use across a wide variety of topics may be appropriate. If the focus is on collecting feedback for program improvement (i.e., formative evaluation—see Rossi, Freeman and Lipsey 2004), then a qualitative or mixed methods approach can provide useful information. Focus groups (Israel and Galindo-Gonzalez 2008), personal interviews, and Sondeos (a type of group interview; see Galindo-Gonzalez and Israel 2008) are some of the methods for collecting qualitative information.

For programs with a medium level of intensity and importance, more time and effort might be warranted. An evaluation for a minor program with multiple activities can measure outcomes at one of the following levels: learning (which includes knowledge, skills, and attitudes), behavioral intent, and behavior change. Because behavior change requires follow-up data collection, more resources are needed than for an end-of-the-meeting (post-only) survey measuring behavioral intent. In addition, more time is needed when an accurate measure of learning or behavior change is desired because surveys with good questions require more than a “top of the head” thought process; they benefit greatly from asking a colleague for a critical review during the development process. Finally, evaluations which collect data from only program participants cannot provide definitive information about impacts. Rather, they are suggestive of the potential impact or when multiple counties are involved, the likely impact. Only when Extension is the sole information source can claims about actual impact be made.

The largest proportion of time and effort spent on evaluation should be focused on one or two of your best programs

(and probably only one at any given time). Because these programs are important, faculty with expertise in evaluation procedures, survey design, and data analysis should be recruited as a partner. “Best” programs warrant having reliable and valid measures to assess learning, behavior change, and other impacts. County faculty can achieve “economies of scale” by partnering with other county faculty, members of statewide teams, and evaluation specialists to collect data from large numbers of program participants and, sometimes, a comparison group of nonparticipants (see

Israel, Easton and Knox 1999; Kropp et al. 2018). Such team efforts can lead to highly credible statements of program impact.

## Guideline #3: Increase/decrease rigor based on available time, money, and expertise

It is important to recognize that evaluations are not free. More intense and rigorous evaluation takes more time, resources, and expertise. One way to save time is by enlisting volunteers to help with specific tasks (e.g., conducting a telephone survey or entering data into a computer file). Including the cost of evaluation activities in grant proposals is a way to increase resources for more rigorous evaluations. As a rule of thumb, 5–10% of the proposal’s budget should be allocated for evaluation activities. On the other hand, plans for evaluations might need revision if the office experiences a budget cut. Finally, having knowledge and skill in designing valid and reliable evaluation tools is a prerequisite for conducting evaluations that are more rigorous. If you have limited expertise in survey design, evaluation procedures or data analysis, then adding a person who has the needed skills to your team can help you conduct more rigorous and useful evaluations.

## Guideline #4: Meet supervisor expectations for quality evaluation data

During meetings with your supervisor, discuss expectations for evaluating your extension programs. Be prepared to negotiate over which programs must be evaluated with a high level of rigor and those which can be evaluated with less rigor (or not at all). Also, review UF/IFAS guidelines for permanent status and promotion at <https://hr.ifas.ufl.edu/tenure-and-promotion> to ensure that your evaluation plans are consistent with University of Florida requirements.

## Guideline #5: Partner with teams and/or other agents to share the work

The adage that “two heads are better than one” also applies to evaluations. Working with a team can bring additional expertise to the table and result in higher quality instruments and procedures. Teams which include a diversity of experience and expertise create evaluation instruments that have better questions and, in turn, less measurement error. The challenge, of course, is to engage all of the team members and to get feedback in a timely manner.

One advantage of working with a team is that administrators are more likely to provide additional resources. These resources can be used to collect data from a larger number of program participants or to gather data from a comparison group of nonparticipants. When county faculty collaborate on an evaluation and collect data on the same outcomes, the data can be combined for the statistical analysis to provide extra power and credibility.

## Guideline #6: Design the evaluation to help yourself

Too often, Extension professionals approach evaluation as a chore, the main purpose of which is to report information to supervisors and external stakeholders. Evaluations can and should be designed to serve your needs, specifically to help you identify when your programs work well and where they can be fine-tuned to provide greater benefits. For example, an evaluation that compares the characteristics of participants who have changed their behavior with those that did not can help identify groups (or “market segments”) where a different approach to teaching might be needed.

One key to a successful career in Extension is being able to document your accomplishments, especially during the permanent status and promotion process (i.e., via your T&P packet). Clearly, it is important to incorporate more rigorous evaluation data for your best programs (see Table 1) to meet increasingly higher expectations of university administrators.

## Guideline #7: Turn lemons into lemonade

It is natural to wish to obtain positive results from an evaluation. In particular, when you are evaluating your extension program you might even feel pressured to show a

positive impact on your clientele. However, the truth is that well-designed and well-implemented evaluations are likely to yield both positive and negative results; it is also true that the latter can be as useful, or even more useful, than the former for program improvement. If after evaluating your extension program you found out that things are not going as well as you thought, you might want to use the results to identify and understand the deficiencies of your program and to determine if something can be done to improve it. It is desirable to involve other stakeholders in designing the improvement strategy to make sure that everybody is comfortable with the proposed changes for the program.

## Summary

Our goal has been to help extension faculty recognize that there is not a “one size fits all” form of evaluation and that it is important to tailor evaluation activities to the situation. The tailoring process involves balancing best evaluation practices with the individual’s and stakeholders’ needs, as well as being cognizant of available resources and implementation intensity to arrive at an appropriate evaluation strategy. For help in setting evaluation priorities and selecting a strategy, contact the UF/IFAS Program Development and Evaluation Center at <https://pdec.ifas.ufl.edu>, or contact one of the authors.

## References

- Galindo-Gonzales, S. and Israel, G.D. 2008. *Using Sondeos for Program Development and Evaluation*. AEC 386. Gainesville: University of Florida Institute of Food and Agricultural Sciences. <https://edis.ifas.ufl.edu/WC067>
- Israel, G.D., Easton, J.O. and Knox G.W. 1999. “Adoption of Landscape Management Practices by Florida Citizens.” *HortTechnology*, 9(2): 262–266
- Israel, G.D., and Galindo-Gonzalez, S. 2008. *Using Focus Group Interviews for Planning or Evaluating Extension Programs*. AEC 387. Gainesville: University of Florida Institute of Food and Agricultural Sciences. <https://edis.ifas.ufl.edu/PD036>.
- Kropp, J., Abarca Orozco, S. J., Israel, G. D., Diehl, D. C., Galindo-Gonzalez, S., Headrick, L. N., & Shelnutt, K. P. 2018. A Plate Waste Evaluation of the Farm to School Program. *Journal of Nutrition Education and Behavior*, 50(4), 332-339.
- Rossi, P.H., Lipsey, M.W., and Freeman, H.W. 2004. *Evaluation: A Systematic Approach*. 7th ed. Newberry Park, CA: Sage Publications.

Table 1.

Importance	Situation	Stakeholders*	Evaluation Strategy	Use & Benefits
Low	Client office visit, phone calls, & e-mails	Self, CED, DED	Occasional customer satisfaction survey	Feedback on service quality; identify ways to improve service to clients
	“One-off” presentations	Self, CED, DED	Occasional customer satisfaction survey or post-only survey	Feedback on teaching and service quality; perceptions of learning; can sharpen presentation skills
	Train-the-trainer workshop	Self, CED, DED	Occasional Train-the-trainer evaluation, which can include satisfaction, learning, and behavioral intentions	Feedback on logistics, learning environment and teaching quality
	Minor program with multiple activities	Self, CED, DED	Self-developed post-only survey of behavioral intent; pre- and post-test or post-then-pre of learning and/or behavior change	Outcome data for ROA and accountability reports; feedback on <i>potential</i> impact
	My “good” program with multiple activities involving other agents (same county)	Self, CED, DED, colleagues	Self-developed pre- and post-test or post-then-pre of learning and/or behavior change	Outcome data for ROA and accountability reports
	My “good” program with multiple activities involving other agents (multiple counties)	Self, CED, DED, colleagues	Self-developed pre- and post-test or post-then-pre of learning and/or behavior change	Larger samples yield more credible data on <i>likely</i> impact for ROA and accountability reports
	My “best” stand-alone program	Self, CED, DED	Consult with evaluation specialist to develop tools and procedures for follow-up evaluations to assess outcomes	Instruments yield more precise measurement of learning and/or behavior change; more credible information in T&P packets and accountability reports
	My “best” program tied to a focus area	Self, CED, DED, Focus team	Coordinate with focus team to develop & use rigorous measures to assess impact	Large samples and comparison groups yield most credible estimates of impact for T&P packets & accountability reports
High				

\*CED and DED refer to UF/IFAS Extension County Director and UF/IFAS Extension District Director, respectively.