

6-28-2019

Situational Complexity and the Perception of Credible Evidence

Scott Chazdon

University of Minnesota, schazdon@umn.edu

Samantha Grant

University of Minnesota, samgrant@umn.edu

Follow this and additional works at: <https://scholarsjunction.msstate.edu/jhse>



Part of the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Chazdon, S., & Grant, S. (2019). Situational Complexity and the Perception of Credible Evidence. *Journal of Human Sciences and Extension*, 7(2), 4. <https://scholarsjunction.msstate.edu/jhse/vol7/iss2/4>

This Original Research is brought to you for free and open access by Scholars Junction. It has been accepted for inclusion in *Journal of Human Sciences and Extension* by an authorized editor of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.

Situational Complexity and the Perception of Credible Evidence

Scott Chazdon

Samantha Grant

University of Minnesota

Situational complexity is the distinction between simple, technically complicated, socially complicated, and complex situations. Programs that operate in simple situations are usually able to follow a prescribed course of action, or recipe, while programs operating in more complicated or complex situations must be flexible and responsive. In this article, the authors present findings from an exploratory, multiple-case study of the credibility of evidence in four distinct program situations ranging from simple to complex. Key informant interviews were conducted with 16 key informants, both internal and external to Extension. The findings were generally that the more complex the situation, the more likely that flexible or mixed-methods approaches were employed to strengthen program credibility. Across all the cases, the relationships that Extension educators have built with stakeholders played a pivotal role in building credibility of evidence. We conclude that sometimes situational complexity matters, sometimes methods matter, sometimes reporting style matters, but what always matters is the trusting relationship between the delivery organization and the stakeholder.

Keywords: complexity, credible evidence, evaluation, Extension education, program stakeholders, qualitative research, comparative case study

"A situation can be described as more or less simple, complicated, or complex. Utility resides in examining the implications and insights generated by asking to what extent a situation is usefully approached as simple, complicated, or complex, or some combination of the three."

—Patton (2011, p. 85)

Introduction

Many of the articles in the *Credible and Actionable Evidence* text (Donaldson, Christie, & Mark, 2015) challenge us to think about the relationship between evaluation methodological choices and credibility. While this relationship is important, it is equally important to think about the relationship between the *situational complexity* of our programs and credibility. Researchers and evaluators have been attuned to the importance of situational and program context for decades

Direct correspondence to Scott Chazdon at schazdon@umn.edu

(Rog, Fitzpatrick, & Conner, 2012; Stake, 1990). Our use of the term *situational complexity* refers specifically to the distinction between simple, technically complicated, socially complicated, and complex situations—a distinction attributed to the work of organizational theorists Ralph Stacey (1996) and David Snowden (2002) and applied more recently to program evaluation by Zimmerman (2001) and Patton (2011).

The situational complexity framework adapted from Stacey's work (Zimmerman, 2011), known as the Agreement & Certainty Matrix (see Figure 1), has two poles based on the level of certainty about cause and effect to solve a problem and the level of agreement among stakeholders about the desirability of the solution. To use a recipe metaphor (Patton, 2011), some Extension programs can actually provide a recipe—a relatively straightforward solution to known problems. These programs typically operate in simple situations where there is a high degree of certainty about cause and effect of a problem and a high level of agreement among stakeholders about the solution. Other Extension programs offer expertise-based frameworks or principles for action in response to more complicated or complex situations where there is more uncertainty about cause and effect of the problem, more social conflict or disagreement about solutions, or both. In these situations, following a recipe does not yield good results.

If program delivery needs to be attuned to situations of varying complexity, becoming less recipe-like as complexity increases, we wondered if evaluation evidence and perceptions of credibility of this evidence would also follow this pattern, becoming less recipe-like as complexity increases. If, as Greene (2015) noted, “method is always the servant of substance, never the master” (p. 206), surely the situational complexity of the program is part of the substance that we cannot ignore. Situational complexity is likely an important influence on credible evidence, so we sought to learn more about how situational complexity mattered.

In this article, we present an overview of situational complexity, and more broadly, program context, and present findings from four case studies examining the credibility of evidence in distinct program situations ranging from simple to complex. We used a comparative case study approach (Ragin & Amoroso, 2011) to learn more about patterns of similarity and difference in the perception of credible evidence. While this was an exploratory type of case study rather than an explanatory, hypothesis testing approach (Yin, 1993), we were curious to see if a credibility pattern would emerge that is similar to the pattern of program delivery, with more recipe-like approaches to evaluation being perceived as more credible for programs operating in simpler situations, while less recipe-like, perhaps mixed-method or participatory approaches would be perceived as more credible for more complicated or complex situations. Understanding the importance of situational complexity for evaluation has implications for how we design and develop future evaluative strategies.

Situational Complexity and Evaluation Design

Before moving to a discussion of situational complexity, it is important to note that evaluation scholars have written about the importance of the related concept of program context more broadly. In fact, an entire issue of *New Directions for Evaluation* focuses on context and its influence on evaluation practice (Rog et al., 2012).

Noting that context matters deeply in evaluation, Rog (2012) identified five areas of context that affect evaluation practice: 1) the context of the problem or phenomenon, 2) the context of the intervention, 3) the broader environment of the intervention, 4) the context of the evaluation itself, and 5) the broader decision-making context (Julnes & Rog, 2015; Rog, 2012). Conner, Fitzpatrick, and Rog (2012) outlined a context assessment process with guiding questions about each of these areas of context that can be used for program planning and evaluation implementation. To clarify what the authors mean by areas of context, some examples of questions used for this context assessment process are shown in Table 1 below.

Table 1. Examples of Context Assessment Questions

| Area of Context | Sample Questions for Context Assessment During Evaluation Planning |
|--|---|
| General phenomenon/problem | What is the problem the program is addressing? What groups prompted concern about it? |
| Particular intervention | Where is the program in its life cycle? Who does the program serve? |
| Broad environment around the intervention | Are there political and social views that affect perspectives on the program, its clients, or decision makers? |
| Parameters of the evaluation | What resources are available to support the evaluation? |
| Broad decision-making arena | What are the main decision makers/users of the evaluation? What are the political expectations for evaluation? |

Note: From Conner et al. (2012), p. 96.

The idea of situational complexity seems closest to “broad environment around the evaluation,” and the idea that the views of program stakeholders surrounding an intervention are highly relevant to the credibility of the evaluation.

Our interest in situational complexity arose from our experience as internal and often developmental evaluators at University of Minnesota Extension. In our roles, we are confronted not only with variation in these five areas of context, but also with the fact that many of the situations in which Extension programs are either newly emerging or constantly changing, often unpredictable, and with a constant need to adapt programming as well as evaluation to these changing situations.

Based on the ideas of organizational theorist Ralph Stacey (Stacey, 1996), Brenda Zimmerman conceptualized a situational matrix to better understand the elements of complexity (Patton, 2011; Zimmerman, 2001). The Agreement & Certainty Matrix plots two features of a situation:

- 1) The level of certainty about cause and effect that can solve a problem, and
- 2) The level of agreement among stakeholders about the desirability of the solution.

In this matrix, there are four program situation types based on these two poles: simple, technically complicated, socially complicated, and complex. The two main distinctions that differentiate programs are the level of agreement and certainty about how to solve the problem. Certainty refers to the predictability about how to solve the problem, and agreement refers to the amount of conflict about how to solve the problem.

Figure 1. Agreement & Certainty Matrix

| | | | |
|------------------|----------|----------------------|-------------------------|
| Agreement | Far from | Socially complicated | Complex |
| | Close to | Simple | Technically complicated |
| | | Close to | Far from |

Certainty

Note: Adapted from Zimmerman, 2001; found in Patton, 2011.

Simple situations are those that are close to certainty about cause and effect of the intervention to solve the problem and are close to agreement among stakeholders on the proposed solution to the problem. These are the recipe-like situations for which “best practices” can be found and agreed upon. Patton shared that in a simple program, the standard procedures that have worked to produce the desired outcomes in the past are highly likely to work again in the future (Patton, 2011). This is not to say that the program is simplistic; rather, it is more formulaic in design. In Extension, there are a variety of best practice programs in which research and evaluation have demonstrated that outcomes can be replicated with fidelity to the program design.

Technically complicated situations are far from certainty about how an intervention will produce a desired effect but are closer to agreement among stakeholders on the proposed solutions (Patton, 2011). Often in these situations, more than one area of expertise is needed, and therefore, must be coordinated and integrated. The solution to the problem is complicated but

knowable. Scaling programs across different program sites is one example of a technically complicated situation.

Socially complicated situations are far from agreement about solutions to the problem but are close to certainty about cause and effect (Patton, 2011). Situations with many different stakeholders offering differing perspectives, articulating competing values, and posing conflicting solutions are socially complicated. Relying on multiple perspectives means that each party has a different take on the situation. Multiple perspectives are important for both innovation and consensus but negotiating interpersonal or intra-group dialogue can be more challenging than actually solving the problem at hand.

Complex situations are both far from certainty and far from agreement. In these cases, high uncertainty about how to produce a desired result can fuel disagreement, and disagreements may intensify and expand the parameters of the uncertainty (Patton, 2011). These types of situations often call for innovative approaches, because there is not an easy—or known—solution to the problem.

For the sake of this article, we did not include the element of chaos that is often referenced in models of complexity, such as the Cynefin framework (Snowden, 2002). Chaos is when there is extreme uncertainty about how to solve the problem and strong conflict among stakeholders. Programs in a state of chaos are not in the right place to be evaluated because these programs are in a state of uncontrollability and unpredictability.

In this article, we have focused on situational complexity from the vantage point of two evaluators within the Extension program. In *Utilization-Focused Evaluation*, Patton (2008) discussed that the choice of an evaluation design is driven by the people involved and their situation. This involves negotiating the intended use of the evaluation for the intended users. Just as a program situation can be simple, complicated, or complex, an evaluation design can vary in its complexity. That is not to say that simple programs only have simple evaluations and complicated programs only have complicated evaluations.

Patton (2011) noted that situational complexity provides a framework to begin to explore evaluation questions. Patton suggested that simple situations lend themselves to logic models where evaluation is based on proving connections between inputs, activities, and outcomes. Evaluation designs in complicated situations should take into consideration the system, stakeholders, and the context of the program and thus lend themselves to more complex questions that investigate the linkages in the system. Developmental evaluation (Patton, 2011) was envisioned as an evaluation approach for complex situations. Evaluation designs for complex situations require the evaluator to stay attuned to changes in processes and outcomes to create feedback loops that continually inform the program and the stakeholders. This type of evaluation would need to be responsive to the program and the environment and takes system influences into consideration at all times.

From our own experiences as long-term evaluators in Extension, we have witnessed that the situational complexity of the program matters greatly for both program design and evaluation. As evaluators, we often start our projects by trying to better understand the program context before designing an evaluation. In each type of program, credibility of evidence is likely achieved through different approaches and methods. To better understand how approaches and methods changed in different contexts, four unique Extension programs were examined through case study interviews.

Research Design

Our methodological choice for this project was a comparative case study. Ragin and Amoroso (2011) described the goal of comparative research is “to elucidate and explain the diversity within a particular set of cases” (pp. 135–136). In choosing cases for analysis, our analytic frame was situational complexity. Our approach was also interpretive, meaning that the goal was to gain a deeper understanding of the ways that situational context affects the perception of program credibility, not hypothesis-testing usually associated with post-positivist approaches to research. As noted by Greene (1994), interpretive approaches to evaluation and research focus on how a program is experienced by various stakeholders.

We identified four programs offered at the University of Minnesota Extension that exemplified the range of Stacey’s four situation types (Zimmerman, 2001), from simple to complex, and focused our analysis on patterns of similarities and differences across these four cases. We intentionally selected programs from all the different program areas within Extension. It is important to note that these programs were fit into the categories retrospectively based upon the evaluators’ knowledge of the program. Our goal was to find one program that best illustrated each situation type.

Although this was an exploratory type of multiple case study (Yin, 1993), we were curious if we would find a pattern that more recipe-like approaches to evaluation would be perceived as more credible for programs operating in more simple situations, while more flexible, and perhaps mixed-method or participatory approaches would be perceived as more credible for more complicated or complex situations.

To examine perceptions of the credibility of Extension program evaluation evidence, we identified and interviewed key program staff for four Extension programs:

- 1) Private Pesticide Applicator Workshops;
- 2) Children, Youth, and Families at Risk (CYFAR);
- 3) Parents Forever; and
- 4) McLeod for Tomorrow.

During our interviews with Extension staff, we asked for names and contact information for one or two key informant external stakeholders, people who Patton (2008) referred to as primary intended users, who would be in a position to reflect on the credibility of evaluation evidence. In some cases, these key informants were programmatic partners, and in some cases, they were funders. Key informants are people who have special insight or expertise about a topic (Patton, 2015, p. 284). The use of key informants in qualitative research arose in ethnographic studies and has continued to be useful in program evaluation when “there is a need to understand motivation, behavior, and perspectives of our customers and partners” (U.S. Agency for International Development, 1996, p. 1). In total, nine key informants internal to Extension and seven key informants external to Extension were interviewed.

Hour-long semi-structured interviews were conducted via phone or Google Hangout. Primary program staff and stakeholders were given interview questions before the interview. This helped the interviewees feel more comfortable and prepared for the discussion.

The selected case study programs are described in more detail in Table 2 and the narrative below. We also collected and analyzed secondary sources of data—program descriptions and program evaluation reports—from each program. Our interview questions for Extension staff and external stakeholders can be found in the Appendix.

Table 2. Selected Case Study Programs and Situation Framing

| Program | What Program Does | What Problem Is Program Addressing? | Brief Description of Situation (Focusing on Certainty/Agreement) | Situational Complexity |
|---|---|--|--|-------------------------------|
| Private Pesticide Applicator Workshops | Trains farmers to safely apply pesticides | Pesticides can cause damage to the person applying them and to the environment. | Close to agreement about the need to address the issue and close to certainty that success will come from educating a target audience. | Simple |
| CYFAR | STEM after-school education program for at-risk youth | Builds a sense of community and science skills in youth. | Close to agreement that youth programs matter but far from certainty about how to fit 4-H programs for the target audience. | Technically Complicated |
| Parents Forever | Parent education for divorcing or separating couples | Promotes resiliency for families transitioning through separation, divorce, and/or custody change. | Close to certainty that this type of intervention will produce desired effects but far from agreement about the need for this type of programming. | Socially Complicated |

| Program | What Program Does | What Problem Is Program Addressing? | Brief Description of Situation (Focusing on Certainty/Agreement) | Situational Complexity |
|----------------------------|--------------------------------|--|--|-------------------------------|
| McLeod for Tomorrow | Community leadership education | Promotes stronger connections among small towns in a county where competition among communities is the norm. | Far from agreement that this type of program is needed and far from certainty that this type of program will produce the desired effect. | Complex |

Private Pesticide Applicator Workshops

A simple program situation is one that has a high level of technical certainty about how to solve a problem as well as a high amount of social agreement about the program approach. Pesticide Applicator Workshops teach both private landowners (e.g., farmers) as well as professionals (e.g., commercial farmers, turf and landscape businesses) about proper pesticide application. For this article and interview, we focused on the training of private audiences. Minnesota has over 17,000 certified private pesticide applicators who require certification to be able to apply restricted use pesticides to their commodity cropland. Certification must happen every three years, and to keep certification, applicators must either attend a recertification training or complete and pass an exam. Approximately 41% of applicators needing recertification attend an Extension workshop annually. An exam needs to be passed if the applicator does not attend an Extension workshop.

The Private Pesticide Applicator Workshops exemplify a simple program as it serves well to manage a problem (incremental innovation of appropriate pesticide application techniques) through educating a specific audience—landowners. The problem addressed by this program is that pesticides can cause damage to the person applying them and to the environment if applied incorrectly. Problems exist with improper use of pesticides, including issues with health and safety, environmental protection, and agricultural pest management.

The Private Pesticide Applicator Workshop training is a long-running, mature Extension program. The benefit of such a mature program is that the program team has established a formulaic model for delivery and testing of the effectiveness. Extension staff deliver training on proper techniques to apply pesticides and work to ensure that research best practices are shared with the field to promote safer pesticide application. Extension staff members create day-long trainings and deliver them around the state. To create the trainings, Extension staff work with a variety of partners to identify the most pressing needs in pesticide application. They then develop content that is informative and engaging for their audiences and meets the requirements for certification.

Although we chose the Private Pesticide Applicator Workshops as a simple program, it could easily be classified as a socially complicated program as the proper mixing and application of pesticides is a technically (and sometimes socially) complicated challenge. The well-established design of creating adult-based pesticide training for a specific target audience was the reason for its selection as a simple program.

Children, Youth, and Families at Risk (CYFAR)

A technically complicated program situation is one in which there is a high level of social agreement about how to solve the problem but a low amount of technical certainty as to the most effective methods. Technically complicated situations require more than one area of expertise that must be coordinated and integrated. Large-scale programs that have multiple local sites are often good examples, as what fits in one context might not fit in another.

The Children, Youth, and Families at Risk program (CYFAR) was our example of a technically complicated program. The CYFAR program is a national grant-funded program to support children, youth, and families at risk. University of Minnesota Extension has held CYFAR grant funding for this project for five years and developed a program model that had multiple years of testing and refinement. Program staff and funders believe that quality positive youth development programs are an important vehicle to get youth excited about science and more interested in science-related fields. CYFAR funding allowed program sites to design the best ways to deliver such programming, as investors were aware that the local context played a pivotal role in the way in which youth were engaged in the program, and they gave programs latitude to make program decisions that would promote positive youth development.

The CYFAR program was designed to ignite Somali middle school-aged youths' interests in learning about Science, Technology, Engineering, and Mathematics (STEM) while preparing them for higher education. Three clubs met in three different locations in Minneapolis, St. Paul, and Eden Prairie. Club participants were youth who experienced educational barriers such as lack of access to resources, high rates of truancy, and disengagement from school. During the school year, the clubs met weekly and were facilitated by two adults. Each club designed its own approach to building science skills based on the needs of the community, youth interest, and staff leader skills. Youth applied their learning to solve practical and scientific engineering problems. Youth also engaged in activities that built leadership skills needed to pursue higher education and careers. Each summer, youth participated in a University of Minnesota campus immersion where they learned about student life, explored academic interests, and identified steps toward college readiness.

Parents Forever

A socially complicated situation involves many different stakeholders offering differing perspectives, articulating competing values, and posing conflicting solutions to a problem. We

chose a divorce education program as an example of a program operating in a socially complicated space. Although there may be technical agreement that providing parenting education to divorcing or separating couples helps them keep their conflicts away from their children, there is not widespread social agreement that this type of program is needed, or about the best approach for providing this type of intervention. Some critics of divorce education argue that the focal point of education efforts should be before marriage occurs in the first place. Other critics of divorce argue that the program should be required of all couples with children, while others think it should remain voluntary. Practicality has played a role in reducing the “dosage” of this program so that it is provided in an online format, and some critics argue that this does not provide as high quality of an intervention as the traditional face-to-face course. Some argue that programs of this type should require a scientific evidence base. There is also a lack of clarity as to whether this approach works for non-white, high conflict, or same-sex couples.

Minnesota Extension’s response to divorce education is the Parents Forever Program, which began in 1994. In 1998, the Minnesota legislature passed a requirement that in all contested custody or parenting time proceedings, parents of a minor child must attend a minimum of eight hours in an orientation and education program. Parents Forever is based on research that suggests factors such as conflict or financial stress can increase the risk of families going through transitions of separation, divorce, and/or changes in custody. The program is designed to encourage parents to pay attention to their own well-being, attend to their children’s development and the parent-child relationship, and improve the co-parenting relationship. Using these three primary mechanisms (well-being, parenting, co-parenting), the curriculum is aimed at increasing resiliency for families transitioning through separation, divorce, and/or custody change.

Parents Forever is highly regarded, often recommended, in counties across the state. Originally, Parents Forever was only a face-to-face program. More recently, Extension developed an online program offering as well. The online offering is popular because it removes transportation and child care barriers, making it easier for parents to attend. As noted by program staff, however, the online offering does not provide the social connections with other parents that have always been reported to be a valuable component of the program.

McLeod for Tomorrow

Complex situations are those in which high uncertainty about how to produce a desired result fuels disagreement, and disagreements intensify and expand the parameters of the uncertainty. Communities in rural counties often find themselves competing with each other for scarce resources, often with a sense that some communities (often the county seat) get all the resources while other communities struggle. At the same time, rural residents often work or shop in communities other than their own, and problems such as workforce issues, economic

development challenges, natural resource concerns, or natural disasters do not begin or end at city limits. We examined McLeod for Tomorrow as an example of a program designed for this type of complex situation.

McLeod for Tomorrow is an example of “county bridging program” (Rasmussen, Armstrong, & Chazdon, 2011). These nine-month cohort programs are designed to strengthen county-wide community by creating “bridging” relationships of communication and understanding. This is often a challenge in counties where local heritage and pride play an important role, where deep-seated rivalries between communities exist, and where long-standing insider groups have controlled the decisions made. Therefore, these programs are explicitly designed to engage new or young residents of diverse backgrounds from the communities across the county.

In McLeod for Tomorrow, a variety of activities are embedded in the program’s design to build trust, mutual respect, commitment, and political awareness among the program participants and communities. To sustain the program, McLeod for Tomorrow became a non-profit 501(c)(3) organization several years ago, with approximately one-third of its budget coming from the county and the remainder coming from donations and grants. A paid, part-time coordinator manages most of these activities, and Extension’s role is to provide the leadership education component.

While it may sound like McLeod for Tomorrow has a recipe that is followed consistently from one year to the next, the mixture of personalities, ages, and cultural backgrounds of program participants varies each year, and the social and economic context is constantly changing. Furthermore, the program is funded in part by the county and must compete for resources with other programs that are less ambiguous in purpose.

Stakeholder Interviews on Credible Evidence

Interviews with Extension staff and external stakeholders focused on the types of evaluation evidence collected, both formative efforts to improve program quality and summative efforts to measure results (Scriven, 1967), and the perceptions held by external stakeholders about the credibility of this evidence. In this section, we highlight key learning from each case about evaluation design and perception of credibility.

Private Pesticide Applicator Workshops

The evaluation approach at this point in the program’s maturity focused primarily on knowledge gains for participants. In this case, the program used a simple evaluation approach anchored around post-session evaluation as well as checks for understanding using Turning Point technology polling during the sessions. Feedback from these evaluations are viewed by the trainers and helped to inform future training efforts. Educators intentionally chose an evaluation focus on educational gains versus public impact (such as environmental impact on water quality),

as they believe that the program has control over the education but not adoption/compliance with proper pesticide use. Moving into an evaluation design that was focused on environmental impact would require control in the design of the study to understand the connection between training from Extension and changes in practice that would ultimately result in environmental shifts.

In addition to the teaching evaluation, a summative impact evaluation study was conducted annually in 2011, 2012, and 2013 of program participants from that year to learn if the trainings were meeting program goals and if any changes should be made to better meet the needs of the audience. Evaluations were mailed to all 1,000 randomly selected participants who had completed the training, and 44% responded.

Results demonstrated that:

- Nearly 73% of workshop participants made at least one pest management decision based on what they learned at the workshop, and 45% made two or more such decisions;
- Farmers found the workshop modules on Personal Protective Equipment to be the most useful in making pest management decisions; and
- 91% of participants planned on attending a workshop again in the future.

Credibility from the side of the program staff was evidenced by positive evaluation data that were fed back into the program to continue to make changes. The program team meets to review all data, and some of the teaching data were also shared with program partners to justify the continued need for the program. The three-year evaluation gave critical feedback to help change content offered. Enrollment for the courses remains high, with an increase of people opting to take the course rather than the exam. In addition, investors continue to give money to the program and are happy with the partnership with Extension.

Perceptions of credibility. Primary stakeholders for this program were investors of the program, including the Minnesota Department of Agriculture, U.S. Environmental Protection Agency, and Extension leaders. Extension staff noted that external stakeholders continue to come back to Extension for this programming, which is strong evidence of the program's credibility as well as trust in the program. They noted that Extension collects summative participant feedback about the program and continues to use formative feedback to improve teaching. One staff member noted that external stakeholders are mostly attuned to "credibility more in how the training is offered or how certain groups are represented."

The external stakeholder interviewed represented a state agency partner who worked closely with University of Minnesota Extension. She expected the Private Pesticide Applicator Workshop training to show results in the form of the summative program outcomes—properly training individuals to adequately apply pesticides without harming themselves or the environment

according to the Minnesota Agriculture Department guidelines. Evaluation approaches, according to her, should also measure the applicability of material and should allow program staff to get feedback about ways to improve the workshop in the future.

This stakeholder was satisfied with the evidence from the program. She trusted the process of collecting the data as well as the quality of data collected. When asked what would cause her to lose confidence in the program, she stated that she would lose confidence if one of three things occurred:

- 1) If there was no evidence that the participants are getting something out of it (no learning occurring),
- 2) If Extension used outdated data to provide recommendations or to adjust the program settings, and
- 3) If there was a mismatch between what the data provides and what was observed in the field.

Interestingly, when asked to think beyond results, the stakeholder noted the important role that Extension staff play in being experts in their field. She stated that she wanted “to see that the University of Minnesota Extension program staff have ample opportunities for professional development and access to the latest research and technologies to stay current with industry trends and maintain credibility from their audience.” This was an example of a stakeholder valuing the people in the organization as much as they value the evidence collected for evaluation purposes.

CYFAR

During the final year of the CYFAR grant, the program team wanted to learn more about the impact of the CYFAR model. Much work had been put into developing a model that could be adapted across program sites, and over the tenure of the grant, the model had been practiced in urban as well as rural environments.

Summative evaluation methods included youth pre- and post-surveys utilizing the common measures provided by the CYFERnet Evaluation Team at The University of Arizona and Virginia Tech. The common measures used in this project measured the impact of 4-H CYFAR program in science and positive youth development. The staff also utilized a formative assessment strategy to document the youth development practices that each site was using each time they engaged with youth. In addition, a mid-year focus group and a final showcase of learning allowed youth to articulate their goals and challenges.

Evidence collected supported the intended program outcomes. Youth liked the program, retention was high, and there was a positive impact on youth’s STEM capabilities. In addition, youth showed increased interest in science careers and STEM abilities. A major thread in the

evaluation was tracking attendance data for youth. In past years of the grant, attendance data were collected by program leaders but not reported back to the principal investigators or evaluators. The team hypothesized that youth with greater participation in the program would see greater science outcome gains. The evaluation did not support this hypothesis, as there were no significant outcome differences for high- and low-participation youth. The majority of youth had high outcome attainment.

The project team remarked that their adherence to mainly quantitative data collection in the form of survey or collection of attendance data did not capture all of the rich stories they heard from program staff. The team decided to conduct both a mid-term focus group and final showcase of learning to better illustrate the benefits of youth programs.

Results have been utilized by program staff in program reports. The team also created videos that have been shared across the 4-H system to document the impact of this program model. Evaluation successes for this group were attributed to the importance of building relationships with program partners as well as with youth. All partners were invested in collecting high-quality data across multiple program sites.

Perceptions of credibility. We interviewed two external stakeholders for the CYFAR program. One was a staff member from a partner non-profit organization that works side-by-side with University of Minnesota Extension to plan, deliver, and evaluate the CYFAR program. In addition, we interviewed a local Minnesota funder of the CYFAR program.

Program staff and partners expressed interest in gathering summative data showing the value of the CYFAR program to scale the program up to other sites. There also was a strong value placed on understanding formative data to inform improvements in the program design that would ultimately increase the impact for youth. This partner stated the evaluation data had been used in meaningful ways for program improvement by stating:

I use the information provided by Extension to improve the program in ways that are in accord with the wants and needs of participants and their parents. We try to shape the program with the vision of participants, and Extension's evaluations are helping us achieve that goal.

The funding stakeholder had great interest in the CYFAR program because it serves a hard to reach an audience of Somali youth. He stated, "We want to help the University and Extension reach communities that they would not have reached otherwise."

When asked about the credibility of the evidence, both stakeholders vouched for its credibility. One stakeholder knew this program had a steady stream of supporting data that had been shared about the program but could not remember specific evaluation findings. This is another example of a time where the trust in both the organization and the reputation of strong work with the right

audience influenced the credibility of the evaluation. The other stakeholder shared that the participatory approach to evaluation added to the credibility of the findings. He shared:

The source of the data, program participants and their parents, adds to the credibility of the evaluation evidence provided by Extension.

When asked under what circumstances they would lose confidence in the program, one respondent stated:

If it becomes a one-way street. It means that the university adopts a top-down approach and imposes [on] us a way of implementing the program. Also, if the structure of the organization changes, shifting from being flexible to being a “one size fits all” type of organization.

Another respondent stated that he would lose confidence if there were an unexplained decrease in the effectiveness of the program or a decrease in the quality of the services.

Parents Forever

The Parents Forever program collects evaluation evidence with participants after each course, whether in-person or online. The evaluation survey includes a series of 14 Likert-scale questions about participant outcomes in the three main focal areas of the curriculum: parent well-being, parent-child relationships, and co-parenting relationships. These items are intended to collect data for both formative and summative purposes. In addition, participants are asked about their likelihood of following through on five specific action steps:

- 1) Adjusting my parenting to better meet the needs of my child(ren),
- 2) Taking steps to improve my support network,
- 3) Identifying my goals,
- 4) Using one of the co-parenting strategies I learned about, and
- 5) Using one of the financial tools I learned about.

Short reports using data visualizations are then prepared annually for each of the eleven sites that provide Parents Forever training. In addition, quarterly evaluation surveys are conducted with facilitators of the program.

In 2005, program leaders published an article in the *Journal of Extension* based on an impact study of the program (Dworkin & Karahan, 2005). The article cited the program’s success in meeting its objectives. Since then, Extension research staff have conducted numerous studies with the goal of testing and documenting the effectiveness of the Parents Forever curriculum in supporting parents through separation, divorce, and/or custody change (Becher et al., 2015; Becher et al., 2018; Cronin, Becher, McCann, McGuire, & Powell, 2017). Program staff are

working on submitting evaluation evidence to various registries to enable the program to received certification as “evidence-based.”

Perceptions of credibility. We interviewed two Extension staff members and two program stakeholders about their views on credible evidence for the Parents Forever program. One stakeholder leads the Community Education department in a rural Minnesota community and provides referrals to divorcing or separating parents who are required by a court order to complete divorce education. The second stakeholder was a family law attorney who helped provided background on legal issues for the development of the curriculum.

The staff members we interviewed noted that facilitators of the program are social workers and are fiercely devoted to this program. They are on the front lines with the parents. One staff member stated, “When you are in a class with parents for eight hours, it is amazing to see parents come in with their arms crossed, disgruntled because they have to be there, but then they leave telling us thank you and ‘I wish I would have taken this class before we decided to divorce.’” The facilitators also are not familiar with evaluation or questions of methodology. Instead, they bring evaluation into their program by continually making program changes based on conversations with participants, intuition, and feedback from other facilitators.

When asked about the type of evaluation evidence she would like to see, one external stakeholder mentioned the rate of people taking a parent education class (versus not taking such a course) in going back to court in the future because they are unable to come to an agreement. This evidence could focus both on dollar savings for the court system as well as reduced conflict. The stakeholder acknowledged, however, that this type of study would be costly and difficult to conduct.

The other external interviewee commented on the evaluation reports that she receives annually. She noted that the reports are difficult to interpret, sometimes using a numbering system for the 14 program outcomes that is hard to follow. She requested that the reports be made more user-friendly.

However, when asked if they believed the evaluation evidence was credible, both interviewees agreed it was. One respondent noted, “I don’t know what else you could ask! I love land grant universities, and I totally trust everything that comes from there.” The other stated, “I have no reason to believe it isn’t credible.” This same respondent went on to state she would only lose confidence in the program

if the program started ignoring experts in the field, whether they be child development people, other legal experts. If the information presented wasn’t based in scholarly or professional knowledge, or if they kept getting evaluations back from parents that the program isn’t helpful and they ignored that. But I don’t think they would do that.

Due to the relationships that have been formed between stakeholders and the Parents Forever program, a high level of negative evidence would be needed to discredit the program.

McLeod for Tomorrow

Evaluation for the McLeod for Tomorrow leadership program involves pre- and post-surveys conducted during the first and last session of the nine-month cohort program. The pre-survey collects data on the organizational involvement of program participants as well as data on their level of civic involvement. These same questions are repeated in the post-program survey. Analysis of the pre-post data shows behavior change as measured by the percentage of participants who actually increase their level of involvement in organizations as well as the percentage who increase their engagement in civic activity. The post-program survey also includes a retrospective pre-post set of questions about the achievement of leadership competencies in the areas of civic engagement, self-efficacy, self-awareness, cross-community knowledge, and shared vision for the future. These competency data are used for formative as well as summative purposes.

In recent years, evaluation staff have begun to collect community impact data using Ripple Effects Mapping and follow-up surveys with program alumni. Ripple Effects Mapping is a participatory group process that engages program participants as well as other community stakeholders, in paired interviews and large group dialogue about the chain of effects produced by a program (see Chazdon, Emery, Hansen, Higgins, & Sero, 2017; Hansen Kollock, Flage, Chazdon, Paine, & Higgins, 2012). The Ripple Mapping session identified several county-wide efforts that had been created by alumni of the McLeod for Tomorrow program. This awareness then led program staff to design a further evaluative study to quantify, in dollar terms, the economic contribution of the program for the county. This study found, overall, that the program returned a value of \$6.40 for every dollar invested (Tuck, Chazdon, Rasmussen, & Bohn, 2019).

Perceptions of credibility. We interviewed two Extension staff members and two external stakeholders. One of the external stakeholders is the coordinator of the non-profit organization that runs the McLeod for Tomorrow program. The other external stakeholder is a county government administrator.

The Extension Leadership educator noted that when she has presented information about the program to the county board, she does not have to work hard to justify the program. They say, “We love this program, and we know it works.” In fact, one of McLeod’s county commissioners went on to participate in the Minnesota Agricultural and Rural Leaders program, another well-known Extension offering. County Commissioners attend McLeod for Tomorrow workshops when they are invited as well as continually provide about one-third of McLeod for Tomorrow’s funding.

The external stakeholder emphasized that Extension's evaluation efforts have helped her communicate with donors and sponsors about the importance of the program. She did note that materials from Extension need to be "easy to consume," meaning that sometimes the tables and information are too complicated, and thus, it is important to keep reports brief and only present the most important information to stakeholders. She also expressed some concern that program participants may rush through the completion of evaluation surveys, which tends to reduce the credibility of the evidence collected.

When asked under what circumstances she might lose confidence in the leadership program, the interviewee spoke specifically about the Extension educator who has always delivered the leadership content, noting that "if she retired, I'd have to build a new confidence in her replacement."

The county administrator noted, "I also tend to trust information that is provided or compiled by reliable resources such as the University of Minnesota Extension office; organizations that have ethics and good practice in place typically care about the type of information presented by those representing their organization." In reaction to the recent economic contribution study report, she noted that "people like to read about people, most of the figures given are numbers—although great, it won't stick in people's minds. Faces and stories stick in people's minds and cause them to take action."

Table 3 summarizes our findings on the credibility of evidence in these four program contexts.

Table 3. Situation Framing and Insights About Credibility

| Program | Situation Framing | Evaluation Design | Insights About Credibility of Evidence |
|-----------------------------|--------------------------|--|--|
| Pesticide Applicator | Simple | End of training participant survey; Follow up survey | Evidence of learning gains is necessary and credible; Properly trained Extension staff are the most important source of credibility. |
| CYFAR | Technically Complicated | Youth pre- and post-survey; Program attendance data; Program lesson plans | Source of data (both parents and youth) adds to credibility; credibility flows from responsiveness of program, not a "pre-canned" approach |
| Parents Forever | Socially Complicated | End of training participant survey; Facilitator survey; Quasi-experimental study | Keep it simple; Maintain trust; if possible, show monetary value (reduced court expenditures) |
| McLeod for Tomorrow | Complex | Pre and post surveys, alumni follow-up; Ripple Effects Mapping; Return on Investment study | Keep it simple; Protect integrity of data collection; Maintain trust; If possible, show monetary value |

Discussion and Implications for Extension Evaluation

This multiple case study project was exploratory. We were interested in patterns of similarity and difference based on situational complexity. We were curious to see if more recipe-like approaches to evaluation were perceived as more credible for programs operating in simpler situations. Similarly, we wanted to see if less recipe-like, perhaps mixed-method or participatory approaches, would be perceived as more credible for more complicated or complex situations.

Our findings were generally that the more complex the situation, the more likely that flexible or mixed-methods approaches were employed to strengthen program credibility. The Private Pesticide Applicator Workshops and Parents Forever programs were able to rely upon simple survey findings to measure learning gains, and this evaluation design was perceived as credible. The other programs also used survey methods but combined these methods with participatory and qualitative methods to better document the richness of the program experience. In the case of the CYFAR program, staff used focus groups, video stories, and a final showcase of learning to strengthen the credibility of evidence. Evaluation of McLeod for Tomorrow has involved both Ripple Effects Mapping and a combined qualitative-quantitative analysis of the economic contributions of the program to the county to strengthen the program's credibility.

While less recipe-like evaluation approaches were used in the more complicated and complex programs in our case study, it was particularly interesting to note that external stakeholders for these more complex programs emphasized that evaluation reporting should "keep it simple."

One of the external stakeholders interviewed for the McLeod for Tomorrow program felt that compelling stories were just as valuable as the quantitative evidence she was presented. This finding from our exploratory study is worth further research. Is there a relationship between "simplicity" of the situational context and credibility of quantitative forms of evidence? Or stated conversely, do stakeholders in complex situational contexts trust narrative more than numbers because they know the numbers cannot tell the whole story?

A limitation of our case study design was that while we set out to select programs that exemplified the four quadrants of the Agreement & Certainty Matrix, we learned through the interviews that programs did not fit these boxes very well. Situations can trend toward "simple," but they are never simple all the time. For example, stakeholders in the Private Pesticide Applicator Workshops could, over time, become more skeptical about pesticide use in general. Patton (2011) noted that simple, complicated, and complex situations are not always easy to distinguish: "There is no complexity thermometer that gives degrees of uncertainty and disagreement on a standardized, all-purpose scale" (p. 95). Programming situations that are close to agreement or close to certainty are increasingly rare. For this reason, it is best to assume that a range of methods and participatory evaluation strategies are worth pursuing to increase credibility.

In addition, we interviewed key informant stakeholders who already had a relationship with Extension programs. This likely biased some of their responses about the credibility of evidence, because these people were largely friends of Extension. However, we believe this limitation did not compromise our findings of this study because the reality of delivering educational programs is that they are driven by stakeholder investment and the stakeholders chosen for the interviews were the primary intended users of the evaluation findings (Bryson, Patton, & Bowman, 2011; Patton, 2008).

While our study showed a pattern of differences based on situational complexity, we also saw an important commonality among the four cases. The cases consistently highlighted the importance of stakeholder trust and Extension's credibility as a delivery organization. Years ago, Weiss and Bucuvalas (1980) found that "trustworthiness" of a research study was a crucial component influencing decisionmakers' likely use of the study findings. But the type of stakeholder trust we heard about was different than the "trustworthiness" of the evaluation research itself. It was more about the trustworthiness of the delivery organization. As noted by Greene (2015), "the credibility of evaluative evidence is not automatically granted via the use of particular empirical methodologies but rather is earned through inclusive, relational, and dialogic processes of interpretation and action that happen on the ground, in context, and in interaction with stakeholders" (p. 206).

In earlier work on the relationship of public value to evaluation, Chazdon and Paine (2014) found that the credibility of the delivery organization, defined as "stakeholder perceptions of the quality of the public program, as well as the reputation of the delivery organization," was a key component in the public value of a program (p. 108). As noted by Chazdon and Paine (2014):

The 4-H program has a long history and is often revered in rural communities. It is likely that some of the public value resulting from this program is derived from this reputation, but this reputation must be carefully safeguarded and cannot always be taken for granted. Moreover, public universities often benefit from a perception that they offer unbiased analyses and reports that hold up to public scrutiny better than analyses or reports produced by for-profit companies. Yet, even prestigious research universities may lose this reputation as a result of a breach of integrity. (p. 109).

Our case study interviews supported this notion. CYFAR had higher sights in their evaluation to demonstrate the value of their program model. The team thought about the different contexts for youth and different stakeholders as well as showing impacts. Staying attuned to stakeholders' needs is a big part of what keeps Extension credible with stakeholders. For both Parents Forever and McLeod for Tomorrow, staying attuned resulted in using data visualization and brief report formats that were easy for external stakeholders to understand. An implication for evaluators is

that maintenance of organizational credibility is a shared responsibility between evaluators and program staff.

The relationships that Extension educators have built with stakeholders play a pivotal role in building credibility of evidence. Mathison (2015) noted that “information becomes evidence through lived experiences, including professional practice. . . . The more context provided for evidence, the better able we are to judge its credibility” (p. 158). Miller (2015) discussed the way that people judge credible information and finds that trust in an individual is a crucial component. In Extension, these trusting relationships proved to be an important part of how credible evidence was assessed by stakeholders.

For that reason, evaluations in Extension should strive, if possible, to use participatory approaches in which both evaluators and program staff are engaged in the design of the evaluation. Our Extension staff have rich relationships built through ongoing work with stakeholders. They understand the needs and wants of their stakeholders, and in our case studies, they were attuned to the evaluation that would help support their program.

Program support and growth, from our interviews with Extension staff, came both from formative evaluation to improve the delivery of the program as well as summative evaluation that positioned the Extension program favorably with stakeholders. Evaluators should use the expertise of program staff to help craft rigorous evaluations that are attuned to both the improvement of the program and stakeholder needs for impact data.

Looking across our four case studies, a common theme was that credibility is as much, or more, about programs and personnel than it is about evaluation rigor. Evaluation is vital, but it is vital because it protects the integrity of the program delivery organization. Regardless of the program, the external stakeholders we interviewed for this study wanted to generally know that Extension programs were evaluating, and they wanted to know that we know our audiences well, but they did not necessarily want to see the evaluations! When they did want to see them, they wanted them to be short and easy to interpret.

We began this study with a concern that too much emphasis has been focused on methodology in discussions of the credibility of evaluation evidence. We thought situational context was also an important influence on the credibility of evidence. Through our case studies, we have learned that sometimes situational complexity matters, sometimes methods matter, sometimes reporting style matters, but what always matters is the trusting relationship between the delivery organization and the stakeholder. As concluded by Greene (2015), “well beyond good method, making meaningful and consequential judgments about the quality and effectiveness of social and educational programs requires engagement, interaction, listening, and caring” (p. 219).

So it seems that the more relevant question for Extension is “what makes the program credible?” rather than “what makes the program *evaluation* credible?” Evaluation is part of what makes a program credible, but it does not stand on its own.

References

- Becher, E. H., Cronin, S., McCann, E., Olson, K. A., Powell, S., & Marczak, M. S. (2015). Parents Forever: Evaluation of an online divorce education program. *Journal of Divorce & Remarriage*, *56*(4), 261–276. doi:10.1080/10502556.2015.1025900
- Becher, E. H., McGuire, J. K., McCann, E. M., Powell, S., Cronin, S. E., & Deenanath, V. (2018). Extension-based divorce education: A quasi-experimental design study of the Parents Forever program. *Journal of Divorce & Remarriage*, *59*(8), 633–652. doi:10.1080/10502556.2018.1466256
- Bryson, J. M., Patton, M. Q., & Bowman, R. A. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and Program Planning*, *34*(1), 1–12. doi:10.1016/j.evalprogplan.2010.07.001
- Chazdon, S., Emery, M., Hansen, D., Higgins, L., & Sero, R. (2017). *A field guide to ripple effects mapping*. Minneapolis, MN: University of Minnesota Libraries Publishing. Retrieved from <https://www.lib.umn.edu/publishing/monographs/program-evaluation-series#Book%202>
- Chazdon, S., & Paine, N. (2014). Evaluating for public value: Clarifying the relationship between public value and program evaluation. *Journal of Human Sciences and Extension*, *2*(2), 100–119.
- Conner, R. F., Fitzpatrick, J. L., & Rog, D. J. (2012). A first step forward: Context assessment. *New Directions for Evaluation*, *135*, 89–105. doi:10.1002/ev.20029
- Cronin, S., Becher, E. H., McCann, E., McGuire, J., & Powell, S. (2017). Relational conflict and outcomes from an online divorce education program. *Evaluation and Program Planning*, *62*, 49–55. doi:10.1016/j.evalprogplan.2017.02.008
- Donaldson, S. I., Christie, C. A., & Mark, M. M. (Eds.). (2015). *Credible and actionable evidence: The foundation for rigorous and influential evaluations* (2nd ed.). Thousand Oaks, CA: Sage.
- Dworkin, J., & Karahan, A. R. (2005). Parents Forever: Evaluation of a divorce education curriculum. *Journal of Extension*, *43*(1), Article 1RIB6. Retrieved from <https://www.joe.org/joe/2005february/rb6.php>
- Greene, J. C. (1994). Qualitative program evaluation: Practice and promise. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 530–544). Thousand Oaks, CA: Sage.
- Greene, J. C. (2015). How evidence earns credibility in evaluation. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *Credible and actionable evidence: The foundation for rigorous and influential evaluations* (2nd ed., pp. 205–220). Thousand Oaks, CA: Sage.

- Hansen Kollock, D., Flage, L., Chazdon, S., Paine, N., & Higgins, L. (2012). Ripple effect mapping: A "radiant" way to capture program impacts. *Journal of Extension*, 50(5), Article 5TOT6. Retrieved from <https://www.joe.org/joe/2012october/tt6.php>
- Julnes, G., & Rog, D. (2015). Actionable evidence in context: Contextual influences on adequacy and appropriateness of method choice in evaluation. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *Credible and actionable evidence: The foundation for rigorous and influential evaluations* (2nd ed., pp. 221–258). Thousand Oaks, CA: Sage.
- Mathison, S. (2015). Seeing is believing: Using images as evidence in evaluation. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *Credible and actionable evidence: The foundation for rigorous and influential evaluations* (2nd ed., pp. 157–176). Thousand Oaks, CA: Sage.
- Miller, R. L. (2015). How people judge the credibility of information. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *Credible and actionable evidence: The foundation for rigorous and influential evaluations* (2nd ed., pp. 39–61). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: Guilford Press.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Thousand Oaks, CA: Sage.
- Ragin, C. C., & Amoroso, L. M. (2011). *Constructing social research: The unity and diversity of method* (2nd ed.). Los Angeles, CA: Sage.
- Rasmussen, C. M., Armstrong, J., & Chazdon, S. (2011). Bridging Brown County: Captivating social capital as a means to community change. *Journal of Leadership Education*, 10(1). Retrieved from https://journalofleadershiped.org/wp-content/uploads/2019/02/10_1_Rasmussen_Armstrong_Chazdon.pdf
- Rog, D. J. (2012). When background becomes foreground: Toward context-sensitive evaluation practice. *New Directions for Evaluation*, 135, 25–40. doi:10.1002/ev.20025
- Rog, D. J., Fitzpatrick, J. L., & Conner, R. F. (Eds.). (2012). *Context: A framework for its influence on evaluation practice*. *New Directions for Evaluation*, 135.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Chicago, IL: Rand McNally.
- Stacey, R. D. (1996). *Complexity and creativity in organizations*. San Francisco, CA: Berrett-Koehler Publishers.
- Stake, R. E. (1990). Situational context as influence on evaluation design and use. *Studies in Educational Evaluation*, 16(2), 231–246. doi:10.1016/S0191-491X(05)80027-6
- Snowden, D. (2002). Complex acts of knowing: Paradox and descriptive self-awareness. *Journal of Knowledge Management*, 6(2), 100–111.

- Tuck, B., Chazdon, S., Rasmussen, C., & Bohn, H. (2019). *Economic value of the McLeod for Tomorrow leadership program, 2017*. St. Paul, MN: University of Minnesota Extension. Retrieved from <https://extension.umn.edu/community-research/mcleod-tomorrow-leadership-program>
- U.S. Agency for International Development. (1996). *Performance monitoring and evaluation tips: Conducting key informant interviews*. Retrieved from https://pdf.usaid.gov/pdf_docs/PNABS541.pdf
- Weiss, C. H., & Bucuvalas, M. J. (1980). Truth tests and utility tests: Decision-makers' frames of reference for social science research. *American Sociological Review*, 45(2), 302–313.
- Yin, R. K. (1993). *Applications of case study research*. Thousand Oaks, CA: Sage.
- Zimmerman, B. (2001). *Ralph Stacey's agreement & certainty matrix*. Toronto, Canada: Schulich School of Business, York University. Retrieved from https://www.betterevaluation.org/en/resources/guide/ralph_staceys_agreement_and_certainty_matrix

Scott Chazdon, Ph.D., is an Evaluation and Research Specialist with the Extension Center for Community Vitality, University of Minnesota. Dr. Chazdon oversees program evaluation for the Center and conducts applied research activities on rural community development topics.

Samantha Grant is Evaluation Director for the University of Minnesota Extension Center for Youth Development. She evaluates the Minnesota 4-H program and focuses on teaching others about how to present data effectively through visual reports and presentations.

Appendix: Interview Protocol

Interviews with Program Staff

1. Describe the program/intervention.
2. Who are the stakeholders in the program? Who is not a stakeholder? Who are trying to influence with your evidence?
3. To what extent is there certainty among various stakeholders about how to solve the problem, to what extent is there agreement/conflict among stakeholders about how to achieve desired outcomes?
4. Describe the evaluation design. What type of evaluation evidence did you collect? Please tell us more about why you chose the evaluation methods you chose?
5. How was the evaluation implemented? What were the results? Did the stakeholders think the evidence was credible? Why did they think it was credible? Or did they want something different? How have you responded?
6. How have the results been used?

Interviews with External Stakeholders

1. Please tell us about your relationship with the Extension program and/or program staff?
2. What results do you expect from the program?
3. What results do you expect Extension to be able to measure?
4. Are you satisfied that Extension's evaluation efforts provide you the evidence you need that the program is achieving intended results?
5. Beyond the results, are there other aspects of the program that you care about? Please explain.
6. Do you believe the evaluation evidence presented by program staff is credible? Why or why not?
7. Under what circumstances would you lose confidence in the program?
8. Is there other evidence you would want to see about the program?