The Foundation Review

Evaluation at Sunset: Considerations When Evaluating a Program as It Concludes

Blair Beadnell
Evaluation Specialists

Holly Carmichael Djang
Evaluation Specialists

Jan Vanslyke
Evaluation Specialists

Barbara Andersen
Santa Barbara Foundation

Follow this and additional works at: https://scholarworks.gvsu.edu/tfr

Part of the Nonprofit Administration and Management Commons, and the Public Affairs, Public Policy and Public Administration Commons

Recommended Citation
Beadnell, B., Carmichael Djang, H., Vanslyke, J., & Andersen, B. (2017). Evaluation at Sunset: Considerations When Evaluating a Program as It Concludes. *The Foundation Review, 9*(1). https://doi.org/10.9707/1944-5660.1349

Evaluation at Sunset: Considerations When Evaluating a Program as It Concludes

Blair Beadnell, Ph.D., Holly Carmichael Djang, M.A., Jan Vanslyke, Ph.D., and Tatiana Masters, Ph.D., Evaluation Specialists; and Barbara Andersen, M.P.A., Santa Barbara Foundation

Introduction

While the benefits of beginning evaluation efforts at program inception are well known, many organizations simply cannot do so. There are a variety of practical reasons for this: funding concerns, lack of capacity, the need to focus energies on program development and implementation, and changing program goals and activities. As a result, many such efforts begin closer to a program's conclusion — they are often termed "sunset evaluations."

The "sunset" descriptor has been used since the 1970s in such phrases as "sunset review" and "sunset evaluation" — public-policy terms referring to an almost always mandated periodic review of a statute, agency, or program to determine whether it should be terminated, continued, or modified. Here, we reframe the phrase "sunset evaluation" to describe a rigorous and useful evaluation that is conducted at or nearing a program's conclusion. Characteristics of these sunset evaluations are that they are voluntary and are intended to provide a road map for other foundations by describing program effects, accomplishments, and lessons learned.

We previously reported findings from an evaluation conducted at the end of a communitywide effort to improve school food sponsored by the Orfalea Foundation. This sunset evaluation of the foundation's School Food Initiative (SFI) showed positive outcomes from the initiative's activities and provided recommendations for organizations interested in engaging in similar efforts (Carmichael Djang, Masters, Vanslyke, & Beadnell, 2016). Because the evaluation was begun as the foundation was spending down and exiting initiatives, it required creative

Key Points

- While the benefits of beginning evaluation efforts at a program's inception are well known, for a variety of reasons many organizations are unable to do so and instead begin these efforts closer to a program's conclusion.

- Previously reported findings from a sunset evaluation of the Orfalea Foundation's School Food Initiative showed positive outcomes of the initiative's activities and provided recommendations for organizations interested in engaging in similar efforts. Because the evaluation was begun as the foundation's activities were winding down, it required creative design approaches.

- This article uses the evaluation of the Orfalea Foundation's initiative to provide a case example of a rigorous and useful sunset evaluation, and discusses other possible extensions of these kinds of methods.

design approaches. This article's goal is to use the foundation's SFI evaluation as a case example showing methods for engaging in this kind of sunset evaluation. This example illustrates approaches we implemented as well as other extensions of the methods used.

Case Example:
A Foundation's Perspective

Philanthropists are increasingly choosing to donate all of their wealth within their lifetime, instead of holding it in perpetuity. As a result, foundations are building timelines for spending down and exiting support for programs. This

was the situation with the Orfalea Foundation, which had invested in the SFI over a long-term period and when the Obama administration elevated school-food reform to a national priority. Foundation leaders wanted to share initiative best practices and lessons learned with other donors and foundations upon their exit.

The SFI is a case in which conducting a sunset evaluation was a useful, even necessary, solution. Over an eight-year period, Orfalea had invested $14.3 million in operating its own educational programs as well as providing grants to schools, school districts, and other nonprofit partners. The goal of these efforts was to improve the quality of food served in schools in Santa Barbara County, California. During this period, the foundation prepared for evaluation by developing a logic model and by having both foundation personnel and grant recipients collect field data.

However, there was minimal capacity to integrate or analyze this data and to prepare comprehensive evaluation reports. Thus, the foundation used these data primarily to make real-time adjustments with a given grant recipient or to provide updates when reporting to the board. For many small to medium-size foundations, using data in this limited way is common. While desirable and beneficial, it is far less common for foundations to integrate evaluation findings across grant recipients in order to draw evaluative conclusions about a complex, multi-year program. There are a variety of reasons for this: foundations may prioritize directing funds to programming rather than to evaluation, or their personnel may have limited evaluation expertise or face competing leadership and organizational priorities. Sunset evaluations, fortunately, can provide organizations that have engaged in limited data-collection and evaluation efforts the opportunity to salvage the data they have collected and better understand and evaluate their overall efforts.

Evaluator Approach in Sunset Evaluation

Before presenting the methods we used in our case example, it is useful to point out two "soft skills" woven through each of the methods. Soft skills are typically defined as behaviors associated with well-functioning relationships with other people, such as communication, interpersonal and social skills, management practices, and leadership. These skills — sometimes understood as emotional intelligence in action — are frequently undervalued in professional settings, but their use can often differentiate between average and outstanding performance (Goleman, 1998; Wilkins, 2014). The two most salient soft skills we identified when reflecting on this case example are collaborative spirit and group process facilitation.

Collaborative Spirit

Intensive collaboration between organization staff and the evaluators was the most essential of the soft skills woven through the SFI evaluation. First, we worked to develop effective relationships with foundation staff to foster successful brainstorming sessions. These sessions were particularly important in efforts to identify previously collected data that could be mined. Second, it allowed the identification and engagement of key individuals (such as front-line staff and community stakeholders) who had knowledge about the context and effects of the initiative. This collaborative process also fostered the buy-in needed to understand the SFI's evolution and identify the effects that had occurred over the previous several years.

Researchers have identified two factors that strengthen the collaborative work of foundation staff and evaluators (Mattessich, Murray-Close, & Monsey, 2001): open and frequent communication, and mutual respect and understanding. Building a collaborative relationship can be as simple as a timely and friendly email response or as complex as tuning into the nonverbal or emotional nuances of communication (Mintzberg, Dougherty, Jorgensen, & Westley, 1996). We used these and other approaches to communication to ensure successful collaboration. One way we demonstrated mutual respect was to explicitly reiterate the importance of foundation personnel input and expertise to this work. We also worked hard to foster a willingness on the part of all parties to step out of their positions

True collaborative spirit involves open and honest dialogue. Engaging in this way in a professional context can be challenging to evaluators and organizational staff alike. Yet this type of emotional intelligence in action is particularly important in sunset evaluations, which require identifying creative solutions for compiling and collecting evaluative information — often in the context of staff reductions, increased demands, diminishing resources, and differing expectations.

as experts in order to hear and learn from one another. Doing so set the norm that all evaluator and staff dialogues were opportunities for mutual learning.

This type of collaboration, however, is more than a set of techniques. True collaborative spirit involves open and honest dialogue. Engaging in this way in a professional context can be challenging to evaluators and organizational staff alike. Yet this type of emotional intelligence in action is particularly important in sunset evaluations, which require identifying creative solutions for compiling and collecting evaluative information — often in the context of staff reductions, increased demands, diminishing resources, and differing expectations.

Facilitating Group Processes

Another key soft skill used in this evaluation was facilitating group activities and discussions to draw out and harness the accumulated knowledge of multiple players. These players were not limited to initiative staff, but also included participants, community members, and other stakeholders. In addition to surveys and individual interviews, we gathered information from various constituencies using facilitated group processes. Group work can help guide participants through a process of assessing their current situation, envisioning and setting goals, developing strategies, and planning action steps. Skilled facilitation of group processes can produce powerful results — for example, an evaluation plan that has been "created, understood, and accepted by all participants" (Wilkinson, 2012, p. 5).

Skillful facilitation of groups is an especially important tool in sunset evaluation. It can maximize the quality and validity of the evaluation's findings in two ways. One way it does this is by eliciting important information that informs the evaluation's design and interpretation. Another way is by providing the benefit of bringing together the inevitable differences in viewpoints that stakeholders have developed over time. For example, we encountered differences in opinion among individuals, all of whom had high investment in the initiative, about the primary outcomes of interest. Maneuvering these differences can be particularly challenging without both access to a range of facilitation techniques and the ability to use them competently. Successful group facilitation by the evaluators or foundation personnel can bring varying perspectives together and position the group to work toward a common viewpoint.

The facilitation techniques we applied in this case were guiding, acting as taskmasters, motivating, and building bridges (Wilkinson, 2012). To do this, we began with a documented facilitators' guide describing how we intended to move the group through this process, though we also allowed for flexibility. By sharing this facilitation guide with the group, we made our plans transparent. This transparency served multiple aims. First, knowing that we had a plan increased

participant comfort. Second, an explicit agenda framed the tasks and helped keep the work on track. Third, seeing progress on these tasks created a great source of energy and momentum, helping us to motivate those involved.

Perhaps the most important facilitation technique we used in this case was bridge building. We encountered many different perspectives on the initiative's theory of change and its evolution. By creating a safe space to discuss them, we were able to illustrate how differences in perspective were useful. This enabled us to find and focus on areas of agreement, and to carefully and respectfully dissect areas of disagreement. We believe these facilitated processes helped foster commitment to the evaluation among foundation staff and board members. In turn, this commitment increased the likelihood that the evaluation would meet their learning needs as well as their desire to share valuable lessons with others.

Methods Useful in Sunset Evaluation
The Orfalea SFI evaluation highlights that while some methods cannot be used at or near program completion, many others are still available. Not usable are methods that must be implemented before a program begins, such as wait-list and randomized-control group trials. However, many other options remain available and appropriate. (See Table 1.) Because sunset evaluation occurs at the conclusion of a program, these options typically involve a process of working backwards. Here, we present three approaches that we used to design and conduct the SFI evaluation: mapping program evolution, leveraging existing data, and collecting retrospective assessments of program effects.

Mapping Program Evolution
Many organizations develop their theories of change and logic models at the beginning of program implementation. These theories and models serve as guides for framing program evaluations because they explain how resources will be dedicated and what effects upon targeted populations are expected. However, many times these frameworks shift during a program's life, whether tacitly or intentionally. In sunset

Key elements in the process of mapping program evolution include understanding the initial theory of program change, factors that led to adjustments, decisions made accordingly, and the concluding theory of change. By engaging in this process the evaluator could learn, for example, that the program shifted its activities specifically because program staff found that a given activity was too cumbersome to implement. This information itself is an important evaluation finding that can help others avoid going down problematic paths when doing similar kinds of work.

evaluations, it therefore becomes important that evaluators understand the history of the program, including how guiding theories evolved over time.

Key elements in the process of mapping program evolution include understanding the initial theory of program change, factors that led to adjustments, decisions made accordingly, and the concluding theory of change. By engaging in this process the evaluator could learn, for example, that the program shifted its activities specifically because program staff found that a given activity was too cumbersome to implement. This

TABLE 1 Approaches to Conducting Sunset Evaluations

Approach	Map Program Evolution	Leverage Existing Data	Retrospective Assessment of Program Effects
Purpose	Understand program evolution: • Initial theory of program change • How and why adjustments occurred • Concluding theory of change • Difference in stakeholder perspectives about theory of change	• Identify outcomes to be measured in any new data collection • Answer evaluation questions	Answer evaluation questions
Potential Data Sources	• Archived documents (e.g., theory of change, logic models, grant applications, internal communications documents) • Current and prior website content • Stakeholder surveys • Staff, stakeholder interviews • Facilitated, structured activities with stakeholders to map understanding of program theory of change • Facilitated discussions with organization leaders about maps	**Primary data:** • Past interview data (formal, informal) • Past survey data • Program participation data • Participant demographics **Secondary data:** • Community surveys • Government records	• Participant questionnaires administered at end of program • Participant interviews conducted at end of program
Outcomes	• Identification of drivers of change • Identification of barriers to, facilitators of implementation • Creation of agreement about evaluation questions to pursue	• Creation of new scales and variables to be included in additional data collections • Assessment of participant characteristics • Evaluation of program effectiveness, overall and for subgroups	Evaluation of program effectiveness, overall and for subgroups
Strength/ Weakness	**Strength:** Evaluators and stakeholders are able to understand program theory of change and shifts in this theory. **Weakness:** Understanding of the program is circumscribed by who participates and what archived data are available.	**Strength:** Cost- and time-effective **Weakness:** Limited by availability, completeness, and reliability of data	**Strength:** Provides findings on participant response to program **Weakness:** • Subject to inaccurate recall • Limited empirical data on validity for CSEPP approach

Tools

information itself is an important evaluation finding that can help others avoid going down problematic paths when doing similar kinds of work. Similarly, it is important to understand if stakeholders had differing perspectives and if so, the degree to which these contributed to refinements in the theory of change. This knowledge can be particularly illuminating. For example, it may turn out that leadership had perspectives on program goals or pathways that differed from those of program staff, or that external stakeholders were not aware of a specific program activity.

There are a variety of approaches in the evaluation literature that describe processes using images or maps to represent program theory and evolution. The two most common are "concept mapping" (Kane & Trochim, 2007) and "outcome mapping" (Earl, Carden, & Smutylo, 2001). While each describes a distinct approach, all share a focus on using participatory methods to bring out diverse perspectives to create insight, understanding, and consensus among stakeholders. The goal and challenge for an evaluator in using these methods, for a sunset or other evaluation, is to facilitate and manage a process that helps participants develop a shared vision of a program's goals, evolution, and outcomes.

In the evaluation of the SFI, we used elements of these approaches to map the initiative evolution by reviewing archived documents and facilitating discussions with organizational leadership. We first reviewed logic models that had been created in each year of the initiative. Doing so allowed us to identify changes across years. We then facilitated structured activities in which stakeholders independently mapped what they believed the theory of change was at the beginning, middle, and end of the initiative, and then compared their thoughts to those of other stakeholders. Following this activity, we facilitated a discussion about the similarities and differences in their maps, with brainstorming about the reasons behind the differences as well as the reasons changes had occurred. In this way, both evaluators and stakeholders were able to understand how the initiative evolved, the drivers behind change, and barriers and facilitators to successful implementation. Additionally, the activity helped

Most programs accumulate data throughout the course of implementing their program, whether or not it is documented. However, program staff may not realize the value of these types of data for use in program evaluation. While it is not unusual to harvest program data for evaluation purposes, this practice is especially useful in sunset evaluations, with the particular advantage of being cost-effective.

Tools

clarify and create agreement about the evaluation questions to be pursued.

While not used in this example, extensions to these approaches exist. For instance, evaluators can uncover programmatic shifts from other types of archived documents, such as mission statements and internal program documentation. Another potential source of archived data is previous content from an organization's website, since that is a location where programs often publish their goals and intentions. Reviewing both current and prior website content can uncover changes in goals and the theories underlying the change process. Using stakeholder surveys is another option for gathering input on past and current program goals as well as shifts in focus that occurred.

Leveraging Existing Data

Most programs accumulate data throughout the course of implementing their program, whether or not it is documented. However, program staff may not realize the value of these types of data

for use in program evaluation. While it is not unusual to harvest program data for evaluation purposes, this practice is especially useful in sunset evaluations, with the particular advantage of being cost-effective (Bamberger, Rugh, Church, & Fort, 2004). Such data could be as simple as the institutional knowledge among program personnel that evaluators formally collect through interviews. Alternatively, it could be data that the program team collected for reasons other than evaluation, such as program participation rates or attendance. Weitzman and Silver (2013) argue for the use of existing data in program-evaluation activities, and point out that while they may not always have all the information desired, they often have information that is useful or closely linked.

The case of Orfalea's SFI evaluation provides an example of the value, as well as potential pitfalls, of mining existing data. One evaluation question was whether schools that more fully participated in the initiative had greater improvements in the intended outcomes. To explore this question, we were able to elicit from the initiative director her existing knowledge about how engaged each school was in the SFI. Based on this information, we created a scale that measured the level of engagement of each school. We used this scale to perform a subgroup analysis in which we compared less- to more-engaged schools. Indeed, we found that more-engaged schools perceived a greater need for the initiative services and had greater improvements in outcomes. While creating this measure of engagement was valuable, it points out a potential danger to keep in mind — specifically, that creating measures from staff recollections carries the danger that the knowledge of outcomes may, without them being aware of it, color their assessments. Evaluators must consider — and take steps to eliminate — such threats to the validity of measures developed in this way.

Many sources of existing data can support a sunset evaluation, and some can even allow for the use of a number of traditional evaluation designs (Shadish, Cook, & Campbell, 2002). For example, data may be available that allow for pre- to post-program tests of participant improvement.

Additionally, data may exist on people who have not received a program (or who have received different services) who can serve as a comparison group. An additional, great example of existing data is when a program team has collected demographic information about participants during the course of its program purely to help with recruitment efforts. Evaluators could use that information to conduct subgroup analyses to identify whether a program works better for some people than for others, thereby giving findings more precision, richness, and nuance.

Typically, other sources of data also exist. For example, it is useful to keep in mind the value of institutional knowledge as a type of existing data. Evaluators can learn about this using thoughtful and reflective interviewing techniques, and this information can inform the evaluation design, process, and interpretation. Additionally, when evaluating programs that intend to make community-level changes, secondary data sources and records (such as community surveys or government records) may provide valuable outcome data. Examples include using arrest data to examine how a program influences violence, emergency room data to explore how a program influences access to health care, or population surveys to track behavioral changes.

Retrospective Assessments of Program Effects

In testing whether a program led to the desired changes, evaluators at program sunset can be limited in two ways. First, they may not have baseline data available to calculate whether change occurred. In such cases, evaluators sometimes collect participant perspectives using retrospective questionnaires. Such questionnaires ask participants to rate the direction and amount of change that occurred as a result of the program. This approach provides some information about possible program effects, but does not solve the second limitation, the lack of a comparison (also known as counterfactual) condition. Specifically, evaluators may not have access to individuals or groups who did not

receive the program and who could serve as a comparison group.

A recent innovation in retrospective data collection has shown promise in this regard. This technique is called "counterfactual as self-estimated by program participants" (CSEPP). With this method, evaluators ask participants to answer outcome questions in two ways: an assessment of themselves currently and how they would be had they not received the program. Evaluators then compare the two answers. In effect, participants serve as their own comparison, providing both program and counterfactual data. Research comparing this method to traditional random-controlled designs indicates that it is a valid way of assessing changes in intentions and attitudes, though further research is need to validate it as a method to assess changes in behavior (Mueller & Gaus, 2015).

In the Orfalea SFI evaluation, we used this method to collect data from cafeteria staff representing how things actually were after the evaluation and a counterfactual comparison — how they would have been without the initiative. The technique provided very useful information supporting the effectiveness of the initiative's efforts. For example, the data showed improvements attributable to the initiative in personnel's professionalism and skills, kitchen equipment, technical assistance, and quality of the food served.

Variations on this approach are also available. While the CSEPP approach emulates a comparison group evaluation design, a slightly different question wording gathers data more like a traditional pretest-posttest design. This retrospective approach asks participants to answer based on how things are at the posttest and also to think back and describe how things were for them before the program (Pratt, McGuigan, & Katzev, 2000). Evaluators have used this method over a much longer period compared to CSEPP, and a larger amount of research on it exists. Studies have found retrospectively-collected, compared to pretest-collected, information to correlate more strongly with objective measures (Bray, Maxwell, & Howard, 1984; Howard, et al., 1979).

Moore and Tananis (2009) point out one issue the retrospective method is meant to address, which is that respondents may not correctly understand and interpret questions before an intervention. For example, participants may overestimate their knowledge on a topic before receiving information about it in a program. Hence, their self-assessment of knowledge would be inaccurately high if asked before the program. After reviewing the literature on retrospective data collection, these authors concluded that there is substantial evidence supporting this concern, with the end result being a less accurate estimation of program effects using pre- and posttesting compared to retrospective measures.

An example of our use of this traditional retrospective approach occurred in a different evaluation project. There, we assessed the effectiveness of training we provided to human-service agency staff on evaluation techniques. To do so, our post-training questionnaires asked them to think back and estimate their skill level before the training. This approach allowed us to compare actual post-training knowledge to their estimates of pre-training knowledge.

These methods can be quite useful in sunset evaluation, although they have a number of important caveats. As with any self-report data, respondents' answers are subject to potential sources of error such as presenting oneself in a positive light, imagining they should illustrate an intervention was effective, or misunderstanding survey questions. While both the CSEPP and retrospective approaches are often the only choice for collecting participant data on initiative effects, continued research is needed to further identify the conditions that maximize the accuracy of information collected using these methods.

Recommendations for Foundations Contemplating Sunset Evaluation

For a foundation, there is significant value in investing in a sunset evaluation at the end of a program or initiative. Doing so can address common goals of foundation leadership. For

Sunset evaluation can also contribute to grant recipients' and other partners' learning. For example, engaging in the evaluation process provides an opportunity for such stakeholders to reflect on their contributions and success (or failure), and this can help improve organizational effectiveness and future partnering with foundations or donors.

example, sunset evaluation can capture and describe the impact of a long-term investment, which helps build a legacy of giving for the foundation and a culture of philanthropy in the communities that they serve. More importantly, providing opportunities for foundations to learn from one other can help the philanthropic sector better understand and invest in critical community needs. Sunset evaluation can also help build a unified vision of multiyear or complex efforts among foundation leadership. Engaging various levels of leadership and program staff in the evaluation process — for example, mapping the evolution of an initiative — can unite leadership around a common vision of initiative success. Sunset evaluation can also contribute to grant recipients' and other partners' learning. For example, engaging in the evaluation process provides an opportunity for such stakeholders to reflect on their contributions and success (or failure), and this can help improve organizational effectiveness and future partnering with foundations or donors.

To increase the likelihood of a sunset evaluation's success, foundation leadership and board members have several important tasks. First, they should engage in a facilitated conversation to establish how highly they prioritize evaluation. Once the decision has been made to dedicate resources to sharing foundation learnings through an evaluation, it is paramount to ensure the level of investment aligns with the prioritization and will adequately support the work. Once this has been achieved, foundation staff must vet and select an evaluator or evaluation firm to guide the work. Important evaluator characteristics to take into consideration include the creative design approaches and soft skills discussed above that are necessary for a successful sunset evaluation. There are other important considerations as well. Does the evaluator align with your organizational values? Do they understand the level of resources that you have available and the implications for the scope of work? Do they communicate with you in a way that helps you understand expectations and feel comfortable asking questions? Lastly, depending on the intended audiences and uses for the evaluation findings, the experience, credentials, and reputation of the evaluation consultant or firm may be important to the perceived credibility of the evaluation.

Conclusions

The evaluation of the Orfalea Foundation's SFI is an example of how rigorous evaluation can occur late in the life of a program, even in its final stages. Specialized techniques such as retrospective assessments of program effects play an important role in these types of sunset evaluations. Soft skills like facilitation are equally key, and their importance should not be undervalued. Together these approaches can produce rigorous, useful evaluations while working within the timing of programs drawing to a close. This is good news given that evaluation is an important element in organizations' missions to address challenging social problems, and that the reality of many programs does not position them to begin their evaluation efforts early.

References

Bamberger, M., Rugh, J., Church, M., & Fort, L. (2004). Shoestring evaluation: Designing impact evaluations under budget, time, and data constraints. *American Journal of Evaluation, 25*(1), 15–37.

Bray, J. H., Maxwell, S. E., & Howard, G. S. (1984). Methods of analysis with response shift bias. *Educational & Psychological Measurement, 44*(4), 781–804.

Carmichael Djang, H., Masters, T., Vanslyke, J., & Beadnell, B. (2016). Key ingredients for school food systems: An evaluation of the Orfalea Foundation's School Food Initiative. *The Foundation Review, 8*(2), 20-38. Retrieved from http://dx.doi.org/10.9707/1944-5660.1295

Earl, S., Carden, F., and Smutylo, T. (2001). *Outcome mapping: Building learning and reflection into development programs.* Ottawa, Canada: International Development Research Centre.

Goleman, D. (1998). *Working with emotional intelligence.* New York: Bantam.

Howard, G. S., Ralph, K. M., Gulanick, N. A., Maxwell, S. E., Nance, D. W., & Gerber, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. Applied *Psychological Measurement, 3,* 1–23.

Kane, M., & Trochim, W. M. K. (2007). *Concept mapping for planning and evaluation.* Thousand Oaks, CA: Sage.

Mattessich, P., Murray-Close, M., & Monsey, B. R. (2001). *Collaboration: What makes it work.* Saint Paul, MN: Amherst H. Wilder Foundation.

Mintzberg, H., Dougherty, D., Jorgensen, J., & Westley, F. (1996). Some surprising things about collaboration: Knowing how people connect makes it work better. *Organizational Dynamics, 25*(1), 30–71.

Moore, D. M., & Tananis, C. A. (2009). Measuring change in a short-term educational program using a retrospective pretest design. *American Journal of Evaluation, 30,* 189–202.

Mueller, C. E., & Gaus, H. (2015). Assessing the performance of the "counterfactual as self-estimated by program particpants": Results from a randomized controlled trial. *American Journal of Evaluation, 36*(1), 7–24.

Pratt, C. C., McGuigan, W. M., & Katzev, A. R. (2000). Measuring program outcomes: Using retrospective pretest methodology. *American Journal of Evaluation, 21*(3), 341–349.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs.* Boston: Houghton Mifflin.

Weitzman, B. C., & Silver, D. (2013). Good evaluation measures: More than their psychometric properties. *American Journal of Evaluation, 34*(1), 115–119.

Wilkins, M. M. (2014, December 31). Signs that you lack emotional intelligence. *Harvard Business Review.* Retrieved from https://hbr.org/2014/12/signs-that-you-lack-emotional-intelligence

Wilkinson, M. (2012). *The secrets of facilitation: The SMART guide to getting results with groups,* (2nd ed.). San Francisco: Jossey-Bass.

Holly Carmichael Djang, M.A., is a principal evaluator at Evaluation Specialists, a public health evaluation firm. Correspondence concerning this article should be addressed to Holly Carmichael Djang, Evaluation Specialists, 7040 Avenida Encinas, Suite 104, Carlsbad, CA 92011 (email: Holly@evaluationspecialists.com).

Blair Beadnell, Ph.D., is a principal evaluator and quantitative specialist at Evaluation Specialists.

Jan Vanslyke, Ph.D., is founder of and a principal evaluator at Evaluation Specialists.

Tatiana Masters, Ph.D., is an evaluator and qualitative specialist at Evaluation Specialists.

Barbara Andersen, M.P.A., chief strategy officer at the Santa Barbara Foundation, was the former director of strategic partnerships at the Orfalea Foundation.

Tools