

Evidence-Based Programming Within Cooperative Extension: How Can We Maintain Program Fidelity While Adapting to Meet Local Needs?

Abstract

In this article, we describe how the recent movement towards evidence-based programming has impacted Extension. We review how the emphasis on implementing such programs with strict fidelity to an underlying program model may be at odds with Extension's strong history of adapting programming to meet the unique needs of children, youth, families, and communities. We describe several techniques that Extension professionals can use to balance program fidelity and adaptability. We suggest that Extension stakeholders may be best served when we tailor certain aspects of interventions without changing the intervention's core components that are responsible for positive outcomes.

Jonathan R. Olson
Research and
Evaluation Scientist
CYFAR Technical
Assistance
Coordinator
jro10@psu.edu

Janet A. Welsh
Research Associate
Professor
Prevention Research
Center
jaw900@psu.edu

Daniel F. Perkins
Professor of Family
and Youth Resiliency
and Policy
Department of
Agricultural
Economics, Sociology
and Education
dfp102@psu.edu

The Pennsylvania
State University
University Park,
Pennsylvania

In recent years, there has been a significant push towards developing and implementing evidence-based programs within Cooperative Extension (Dunifon, Duttweiler, Pillemer, Tobias, & Trochim, 2004; Fetsch, MacPhee, & Boyer, 2012; Hill & Parker, 2005; Perkins, Chilenski, Olson, Mincemoyer, & Spoth, 2014). Despite some inconsistencies in definitions of evidence-based programming, most agree that such initiatives have been built upon a sound theoretical and/or empirical base, and their effectiveness has been demonstrated through high-quality outcome evaluations (Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Clearinghouse for Military Family Readiness, n.d.; Elliott & Mihalik, 2004; Flay et al., 2005; Olson, 2010; Small, Cooney, & O'Connor, 2009). The roots of this type of programming are based in the evidence-based medicine movement that began to take hold in the 1990s (Claridge & Fabian, 2005), and has recently extended to include a variety of program areas, including child-, youth-, and family-focused programs that are commonly implemented by Extension professionals.

Scholars have argued that using evidence-based programs comes with a variety of benefits for Extension professionals. For example, evidence-based programming can help increase the overall effectiveness of outreach efforts, can help educators and specialists become more accountable to

funding agencies, and can help demonstrate positive outcomes to various stakeholders (Dunifon et al., 2004; Fetsch et al., 2012; Hill & Parker, 2005). Furthermore, the increasing availability of established, pre-packaged evidence-based programs may encourage more efficient programming, eliminating the need to "reinvent the wheel" when developing new programs (Fetsch et al., 2012; Olson, 2010). In light of these benefits, a variety of scholars have called on Extension administrators, faculty, and staff to increase their commitment to implementing such programs (Dunifon et al., 2004; Fetsch et al., 2012; Hill & Parker, 2005).

How Have Evidence-Based Programs Fared in the "Real World?"

As the number of evidence-based programs continues to grow, we are gaining more information about their effectiveness, as well as factors that either promote or inhibit programmatic success. Recent evaluations of such programs have revealed somewhat mixed findings. While some evidence-based programs have demonstrated effectiveness through high-quality evaluations, several reviews of the literature have suggested that a variety of programs commonly marketed as "evidence-based" have demonstrated only modest or even negligible positive effects when tested in real-world settings (Elliott & Mihalik, 2004; Fetsch et al., 2012; Gandhi, Murphy-Graham, Petrosino, Chrismer, & Weiss, 2007; Olson, 2010). In light of such information, several scholars have proposed potential explanations for why highly regarded programs may demonstrate limited effectiveness when implemented in common community-based settings. The following section of this paper reviews several common explanations for program failure.

Why Might Evidence-Based Programs Fail in "Real World" Settings?

One reason that evidence-based programs might fail within "real world" settings is that they may simply not work for the general public. We would not expect a program to work if it was based on a flawed or incomplete theoretical and/or empirical foundation (Chen, 1990, 1998), or if program activities were not adequately aligned with the program's underlying theory. In short, it is possible that despite positive outcomes in preliminary studies, some programs simply do not adequately address factors that promote behavior change among program participants (Gandhi et al., 2007).

A second possible explanation for observed program ineffectiveness is a flawed evaluation. Common limitations in program evaluations include, but are not limited to, biased samples, poor measurement procedures, high levels of participant attrition, and inappropriate statistical tests (Cook & Campbell, 1979). If an evaluation of a particular youth- or family-focused program has one or more of these flaws, it becomes difficult to determine if a program was truly ineffective, or if the poor evaluation design made it appear that way (Gandhi et al., 2007).

A final common explanation for program ineffectiveness is related to program implementation, which refers to the degree to which all parts of a program are administered as expected. If a program was not properly implemented, it is possible that key components of the program were not fully delivered, rendering the program ineffective (Bumbarger & Perkins, 2008; Chen, 1990, 1998; McHugh, Murray, & Barlow, 2009). For example, if an Extension educator implemented an afterschool program for

adolescents, but due to time constraints eliminated three of 10 lessons, we might see reduced effectiveness as a result of the missing content. Similarly, we may see poorer outcomes if program participants differ significantly from the population for which the program was originally intended. Scholars refer to the degree to which a program is implemented as intended as "program fidelity." When fidelity is high, few changes have been made to the program. When fidelity is low, however, significant changes have been in terms of program content, timing, and/or populations served (Fixsen, Blase, Naoom, & Wallace, 2009; Rossi, Lipsey, & Freeman, 2004).

What Do These Factors Mean for Extension Programming?

Each of the above-mentioned reasons for program ineffectiveness has implications for Extension professionals interested in evidence-based programming. The first two encourage us to carefully consider the quality of program evaluations and to only choose those programs that have demonstrated positive outcomes through multiple high-quality evaluations (e.g., randomized control trials or rigorous quasi experimental designs). Similarly, for those developing their own programs, this means closely monitoring program effectiveness using high-quality evaluations methods.

The third reason has received considerable attention in recent years, likely due to the fact that many practitioners do make changes when implementing evidence-based programs (Bumbarger & Perkins, 2008). Common changes include editing curriculum content to be more culturally relevant or age-appropriate, cutting the length of sessions to fit within time constraints, using the program with a new population for which we have no evidence of effectiveness, and changing lessons to better mesh with other concurrent programs and/or educational strategies (Barrera, Castro, Strycker, & Toobert, 2013; Bumbarger & Perkins, 2008; Castro, Barrera, & Martinez, 2004; McHugh et al., 2009). Because many scholars view program fidelity as fundamental to program success, significant attention has been given to helping program implementers deliver programs in ways that align with program developers' original intent (Durlak & DuPre, 2008; Elliott & Mihalik, 2004; Greenberg, Domitrovich, Graczyk, & Zins, 2005; McHugh et al., 2009).

Does Encouraging Program Fidelity Make Us Less Responsive to Community Need?

A hallmark of the Cooperative Extension System has been its ability to be responsive to the needs of the communities it serves. Thus, some Extension professionals report that they have resisted implementing pre-packaged evidence-based programs that place a strong emphasis on fidelity. Indeed, in their study of Family Living and 4-H Youth Development Educators, Hill and Parker (2005) found that participants viewed traditional Extension programs as being at least as effective as pre-packaged evidence-based prevention programs. Similarly, Fetsch and colleagues (2012) note that community-based Extension professionals oftentimes need to quickly and efficiently react to emerging issues that are unique to a particular region. In such cases, flexibility in programming is important and oftentimes trumps fidelity to an established program model.

Despite these barriers, the evidence-based movement continues to grow. In light of the current emphasis placed on accountability and wise investment of limited resources, Extension personnel need to consider how evidence-based programming can be incorporated into their existing work. In recent

years, a compromise between program fidelity and adaptability has been developing. Despite strong calls for strict adherence to program protocols among many prevention scientists (Elliott & Mihalik, 2004), others are beginning to soften their stance. Indeed, a small but growing group of scholars has begun focusing on how practitioners can balance the push towards fidelity with the desire to be responsive to the unique strengths and needs of individual program participants (Castro et al., 2004; Greenberg et al., 2005; McHugh et al., 2009).

How Can We Balance Fidelity and Adaptation?

Perhaps the most important step that Extension professionals can take when seeking to modify an evidence-based program is to identify the program's core components. These components are activities, practices, and/or lessons that have been found to be responsible for a program's overall effectiveness. They are commonly referred to as a program's "essential ingredients." Core components are directly related to a program's theory of change (Blase & Fixsen, 2013; Fixsen et al., 2009; O'Connor, Small, & Cooney, 2007; Backer, 2001). As such, eliminating core components should be avoided because doing so means eliminating a key ingredient for programmatic success.

However, some program components are flexible in that they can be adapted to better fit with a particular culture or context (Blase & Fixsen, 2013; O'Connor et al., 2007). For example, a program might include optional team-building exercises, discussion sessions, or social events that are not directly related to the theory of change. Such components would be good candidates to be modified, adapted, or eliminated as a way to better meet local needs. In the authors' own work, the PATHS program has specific children's books associated with teaching emotions; however, when we worked with program implementers in Northern Ireland, a culturally relevant children's book that addressed the same emotions was substituted for the original.

Identifying Core Components

Probably the easiest method to identify core components is to contact the person who originally designed the program. Program developers may be able to provide information about the theory upon which the program has been based and whether there have been any previous examples of successful adaptations to the program. In some cases, program designers and/or program evaluators have detailed information on how specific components are related to outcomes. For example, some programs, such as Multisystemic Therapy (MST), have been so extensively studied that we now have a good sense of which components must be implemented. In fact, MST has specific structures in place to guide replications of the program in new settings. Unfortunately, only very few program developers have such detailed knowledge about which components are most strongly tied to desired outcomes (Blase & Fixsen, 2013; Elliott & Mihalik, 2004).

In cases where program developers do not have detailed information about the core components of their programs or in cases where Extension personnel are developing and adapting their own evidence-based programs, additional work is necessary before any adaptations should be made. A good place to start this process is to identify the theory of change upon which a particular program is based. The theory of change for a particular program might be outlined in a logic model. Logic models summarize key program activities, inputs, outputs, and expected outcomes. However, most do not

describe the causal pathways through which program activities eventually lead to the expected outcomes. In short, they usually focus more on what a program does than how it does it (Chen, 1990, 1998; Taylor-Powell, Jones, & Henert, 2003).

In contrast, a program's theory of change focuses on the conceptual framework upon which a program is based. In logic model terms, it focuses on the causal chain that leads from inputs and outputs to program outcomes. Ideally, an explicitly stated theory of change is developed prior to program development, and each program component will have been specifically designed to promote specific outcomes. Such a process makes it easier for program evaluators to determine the extent to which each particular aspect of a program has intended or unintended effects. Indeed, a well-designed theory-driven outcome evaluation can provide a wealth of information about not only if a program worked, but which aspects were more or less effective in impacting outcomes (Bickman, 1987, 1990; Chen, 1990, 1998; Chen & Rossi, 1992; Weiss, 1995, 1997). This process is particularly important if Extension professionals are developing their own "homegrown" programs, as it can help develop an evidence base that is often lacking among such strategies.

In practice, few programs meet the above criteria. Oftentimes, a program is based on a theory of change, but seldom is it explicitly stated with enough detail to tie individual program components to expected or observed program outcomes (Elliott & Mihalik, 2004). As such, we often have a general idea of what the core components may be, but there is typically a fair amount of uncertainty regarding key ingredients for success. This leaves program implementers with an important decision. Without knowing the core components of a program, they can implement the program with complete fidelity to the original model, which ensures that every program component is implemented, or they can collect original data to identify core program components.

Theory-Driven Evaluation and Usability Testing

During the 1990s, Chen and other evaluation experts (Bickman, 1987, 1990; Chen, 1990, 1998; Chen & Rossi, 1992; Weiss, 1995, 1997) argued in favor of theory-driven outcome evaluations. According to these scholars, the purpose of such evaluations is to determine not just if a program works, but how and why it either does or does not produce expected outcomes. To complete this type of evaluation, one must identify a program's underlying theory of change. Next, an evaluation can be designed that assesses the various aspects of that theory. For example, if an after-school mentoring program has been designed to promote academic success by facilitating social bonds among mentors and program participants, an evaluator will want to assess those social bonds along with measures of academic success. If program participation can be linked to increasing bonds among mentors and participants and these bonds are related to later academic success, we can reasonably conclude that the mentoring bond is a core component that facilitates program success.

In some cases, theory-driven evaluations of pre-packaged programs already exist. Indeed, as a result of such evaluations, we currently have a good sense of the core components of several comprehensive programs, such as the MST program described above. However, most pre-packaged programs have not been subject to theory-driven evaluation, and few program developers incorporate theory-driven evaluation designs when they roll out new initiatives (Elliott & Mihalik, 2004). This issue may be particularly salient for Extension professionals who have developed their own programs, because they

would be responsible for developing their own theory-driven evaluations as one way to systematically examine the core components of their programs.

When theory-driven evaluations are either not available or not feasible, Extension personnel might consider using a series of smaller pilot tests to assess the effects of small, incremental changes to a program. Blase and Fixsen (2013) suggest usability testing as a way to determine the effects of changing program components. Usability testing is a form of program evaluation in which a researcher assesses program effectiveness in stages with small samples. This type of testing typically begins with an assessment of program outcomes for a small number of participants who have received the full program with complete fidelity to the original model. Based on the data gathered, small changes can be made to the program, and this new version is then implemented and tested on another small sample. This process can be repeated several times until an effective, modified version of the original program has been developed (Barrera et al., 2013; Blase & Fixsen, 2013).

Applying Usability Testing to Extension Programs

In recent years, scholars have begun to incorporate concepts from usability testing when adapting youth- and family-focused programs. For example, Barrera and colleagues (2013) synthesized a variety of recommendations on how to approach program adaptation in a systematic way. They specifically suggest the following stages:

1. Gather information to ensure that adaptation is justified and which components are most amenable to change,
2. Make preliminary changes to a program's curriculum,
3. Pilot test the effects of the changes,
4. Refine the adaptation, and
5. Conduct a theory-driven empirical trial of the refined product.

To date, there are few examples of comprehensive efforts to adapt programs using such systematic approaches. Unfortunately, most curriculum adaptations tend to be somewhat haphazard and based more on good intentions or pragmatic pressures than sound science. Bumbarger and Perkins (2008) suggest that many modifications to youth- and family-focused programs represent "program drift" in which variations in program delivery result from reactions to barriers rather than planful, theory-based innovation. Common barriers include time constraints, limited resources, and characteristics of the intended audience. In addition, stakeholder preferences might influence programming choices. Sometimes, these changes reflect simple additions or program enhancements that do not affect the core components of the program. In theory, such changes should not impact program effectiveness. However, it is important for program implementers to work with stakeholders to make sure that core program components are not altered or eliminated.

Kumpfer and colleagues have noted how program modifications can have detrimental effects on program outcomes. Despite following a systematic approach to adapting the Strengthening Families

Program, they found that some modified versions of the program were not as effective as the original version (Kumpfer, Alvarado, Smith, & Bellamy, 2002). The authors attributed such findings to a watering down of core program components. Such conclusions have practical implications for Extension professionals in that program implementers could consider adopting previously evaluated variants of established programs, but only if they show effectiveness in high quality outcome evaluations. This approach could save time relative to full usability testing as long as detailed records were kept regarding the nature and extent to which modifications were made to the original program.

Conclusions and Implications for Extension Personnel

As outlined in the beginning of this article, the general movement toward evidence-based programming has been associated with a variety of benefits, and there has been a growing number of calls for Extension to become more evidence-based (e.g., Dunifon et al., 2004; Fetsch et al., 2012; Hill & Parker, 2005). However, strict adherence to evidence-based curricular standards has resulted in a loss of flexibility in program implementation. As described in this article, Extension personnel may be able to find a reasonable balance between the strict emphasis placed on program fidelity and Extension's fundamental strength of being responsive to individual community needs. Fortunately, as outlined in this article, a growing body of literature has identified strategies that can help program implementers balance fidelity and program adaptation. Based on the work of Blase and Fixsen (2013) and other literature reviewed throughout this article, Extension personnel may be most successful by following these steps:

1. Clearly describe the context in which a program will operate:
 - a. Who does the program serve?
 - b. Where will it be implemented?
2. Identify core program components:
 - a. What are the active ingredients?
 - b. Contact program developers and review the literature.
3. Make changes to the curriculum:
 - a. Ensure that core components remain.
 - b. Determine what kinds of changes stakeholders want.
 - c. Discuss with stakeholders the problems with core component adaptations.
 - d. Tailor non-essential components to meet unique needs of program participants.
4. Evaluate the results:

- a. Pilot test the program on a small group of participants following small incremental changes to a curriculum.
- b. Incorporate feedback from the pilot tests into each iteration of the program.

By following these steps, Extension professionals may be able to experience the benefits of evidence-based programming without losing the ability to adapt programs to respond to unique community needs. With staff geographically dispersed throughout the United States, the Cooperative Extension System is uniquely positioned to appreciate the diversity of young people and their families and communities. By striking a balance between fidelity and adaptability, we can tailor our services to meet the needs of all our stakeholders.

References

- Backer, T. E. (2001). *Finding the balance—Program fidelity and adaptation in substance abuse prevention: A state-of-the art review*. Center for Substance Abuse Prevention, Rockville, MD.
- Barrera, M. B., Castro, F., Strycker, L. A., & Toobert, D. J.. (2013). Cultural adaptations of health interventions: A progress report. *Journal of Consulting and Clinical Psychology*, 81, 196–205.
- Bickman, L. (Ed.). (1987). *Using program theory in evaluation*. San Francisco: Jossey-Bass.
- Bickman, L. (Ed.). (1990). *Advances in program theory*. San Francisco: Jossey-Bass.
- Blase, K., & Fixsen, D. (2013). Core intervention components: Identifying and operationalizing what makes programs work. ASPE Research Brief. US Department of Health and Human Services.
- Bumbarger, B., & Perkins, D. (2008). After randomised trials: issues related to dissemination of evidence-based interventions. *Journal of Children's Services*, 3(2), 55-64.
- Castro, F. G., Barerra, M., & Martinez, C. R. (2004). The cultural adaptation of preventive interventions: Resolving tensions between fidelity and fit. *Prevention Science*, 5, 41–45.
- Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *The Annals of the American Academy*, 591, 98-124.
- Chen, H. T. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chen, H. (1998). Theory-driven evaluations. *Advances in Educational Productivity*, 7, 15–34.
- Chen, H. T., & Rossi, P. H. (Eds.). (1992). *Using theory to improve program and policy evaluation*. Westport, CT: Greenwood.
- Claridge, J. A., & Fabian, T. C. (2005). History and development of evidence-based medicine. *World Journal of Surgery*, 29(5), 547-553.

Clearinghouse for Military Family Readiness (n.d.). Programs. Retrieved from:

<http://www.militaryfamilies.psu.edu/programs>

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field setting*. MA: Houghton Mifflin.

Dunifon, R., Duttweiler, M., Pillemer, K., Tobias, D., & Trochim, W. M. K. (2004). Evidence-based Extension. *Journal of Extension* [On-line], 42(2) Article 2FEA2. Available at: <http://www.joe.org/joe/2004april/a2.php>

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 4, 327–350.

Elliott, D. S., & Mihalik, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science*, 5, 47–53.

Fetsch, R. J., MacPhee, D., & Boyer, L. K. (2012). Evidence-based programming: What is a process an Extension agent can use to evaluate a program's effectiveness? *Journal of Extension* [On-line], 50(5) Article 5FEA2. Available at: <http://www.joe.org/joe/2012october/a2.php>

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, 19(5), 531-540.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S.,...Ji, P. (2005). Standards of evidence: criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 6, 151-175.

Gandhi, A. G., Murphy-Graham, E., Petrosino, A., Chrismer, S. S., & Weiss, C. H. (2007). The Devil Is in the Details Examining the Evidence for "Proven" School-Based Drug Abuse Prevention Programs. *Evaluation Review*, 31, 43-74.

Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2005). The study of implementation in school-based preventive interventions: Theory, research, and practice. *Promotion of Mental Health and Prevention of Mental and Behavioral Disorders 2005 Series V3*.

Hill, L. G., & Parker, L. A. (2005). Extension as a delivery system for prevention programming: Capacity, barriers, and opportunities. *Journal of Extension* [On-line], 43(1) Article 1FEA1. Available at <http://www.joe.org/joe/2005february/a1.php>

Kumpfer, K. L., Alvarado, R., Smith, P., & Bellamy, N. (2002). Cultural sensitivity and adaptation in family based prevention interventions. *Prevention Science*, 3, 241–246.

McHugh, R. K., Murray, H. W., & Barlow, D. H. (2009). Balancing fidelity and adaptation in the dissemination of empirically-supported treatments: The promise of transdiagnostic interventions. *Behaviour Research and Therapy*, 47(11), 946-953.

O'Connor, C., Small, S. A., & Cooney, S. M. (2007). Program fidelity and adaptation: Meeting local needs without compromising program effectiveness. *What Works, Wisconsin Research to Practice*

Series, 4, 1-6.

Olson, J. R. (2010). Choosing effective youth-focused prevention strategies: A practical guide for applied family professionals. *Family Relations, 59*, 207-220.

Perkins, D. F., Chilenski, S. M., Olson, J. R., Mincemoyer, C. C., & Spoth, R. (2014). Knowledge, attitudes, and commitment towards evidence-based prevention programs: Differences across Family and Consumer Sciences and 4-H Youth Development Educators. *Journal of Extension (On-line), 52*(3) Article 3FEA6. Available at <http://www.joe.org/joe/2014june/a6.php>

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

Small, S. A., Cooney, S. M., & O'Connor, C. (2009). Evidence-informed program improvement: Using principles of effectiveness to enhance the quality and impact of family-based prevention programs. *Family Relations, 58*, 1-13.

Taylor-Powell, E., Jones, L., & Henert, E. (2003). *Enhancing program performance with logic models*, University of Wisconsin-Extension, Feb. 2003.

Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. Connell, A. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives* (pp. 65–92). New York: Aspen Institute.

Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review, 21*, 501–524.

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

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