Published in final edited form as:

Health Educ (Lond). 2013 ; 113(4): 264–280.

Component analysis of adherence in a family intervention

Laura G. Hill and

Department of Human Development, Washington State University, Pullman, Washington, USA

Robert W. Owens

Department of Educational Leadership and Counseling Psychology, Washington State University, Pullman, Washington, USA

Abstract

Purpose—Most studies of adherence use a single global measure to examine the relation of adherence to outcomes. These studies inform us about effects of overall implementation but not about importance of specific program elements. Previous research on the Strengthening Families Program 10–14 has shown that outcomes were unrelated to global adherence. The purpose of the present study was to determine whether adherence to specific components of SFP was related to outcomes, even though global adherence was not.

Design/methodology/approach—The authors micro-coded data from an observational study of 11 instances of SFP ($N = 47$ facilitators, 151 participants) into specific process and content components. Using multilevel analysis, they examined the relation of each component to program outcomes, accounting for individual- and program-level variables.

Findings—Most associations of adherence with outcome were negligible for European-Americans but significant for minority participants.

Research limitations/implications—Global assessments of implementation are insufficient for complex, multi-component prevention programs and may obscure relations of implementation to outcomes. Additionally, program components may function differently based on participant characteristics.

Practical implications—Facilitators would benefit from understanding the function of individual program components, particularly when programs are delivered to diverse audiences. Program developers should provide detailed logic models of program theory to guide facilitators' decisions about adaptation.

Originality/value—This is one of only a few studies to examine the relation of adherence to specific intervention components to outcomes in a real-world setting. Results show the utility of component analysis and the importance of considering individual characteristics for implementation assessment.

Keywords

Prevention; Parenting intervention; Fidelity; Program implementation; Component analysis; Adherence; Parents; Family

Introduction

Adherence, or the degree to which program components are delivered as prescribed, is an important aspect of the multidimensional construct of program fidelity. In the context of a controlled research environment, adherence to curriculum is achieved through close monitoring of facilitator training and program implementation. However, when efficacious programs are translated into real-world settings, implementation fidelity may vary greatly across sites, and changes in delivery from the program-as-intended may result in lower program effectiveness (Blakely *et al.*, 1987; Kalafat *et al.*, 2007; McDonnell *et al.*, 2007). For this reason, standardization of program content and training protocols is often a precondition for programs to be designated “model” or “exemplary” (Substance Abuse and Mental Health Services Administration (SAMHSA), 2012), and strict adherence to program content is emphasized in trainings.

The overarching goal of the present study was to determine whether assessing adherence to individual program components yielded information about ingredients critical to positive outcomes. When adherence is measured, studies generally report global assessments of how much or how well program content was implemented – for example, adherence may be calculated as the percentage of prescribed program elements that were actually delivered in an implementation. Although global fidelity assessment provides a general picture of overall implementation quality, it does not provide information regarding the importance of specific program elements. In the present study, an extension of a previous observational study of an evidence-based family intervention (Cantu *et al.*, 2010), we micro-coded all elements of a program curriculum, categorized them according to a coding system that separated content components from process components, and examined the association of each component with short-term program outcomes. Our component analysis extends current literature on program fidelity, because there are few component analyses of programs as they are translated into real-world settings. In addition, the analysis contributes to the body of studies that have examined the association of adherence to program outcomes for the specific intervention we observed (the Strengthening Families Program (SFP) for Parents and Youth 10–14). However, it differs from those studies, which reported results of global adherence assessments. In the current study, we demonstrate that component analysis can increase the utility of fidelity assessment by providing information about which ingredients of a program are most closely tied to intervention success (Orwin, 2000).

The emphasis on adherence to content in delivery of evidence-based programs arises from concern for preserving the causal link between intervention and outcome that was established in efficacy trials. Numerous studies have shown that as programs move into real-world settings, adherence tends to be lower than in research trials, and program effect sizes are correspondingly diminished (Durlak and DuPre, 2008). However, other studies have

shown that high adherence does not necessarily lead to expected outcomes when there are structural and contextual barriers to program effectiveness (e.g. Gottfredson *et al.*, 2006). The study of adherence effects is further complicated by concerns about generalizability as programs move into community settings (Fox *et al.*, 2004; Sanders *et al.*, 2002). For example, some studies have shown that adaptation of programs to reflect cultural norms and values may positively affect recruitment and retention of minority participants (Orrell-Valente *et al.*, 1999; Yancey *et al.*, 2006), increasing dosage and thus ultimately increasing effectiveness. However, despite a general move toward cultural adaptation of evidence-based programs (Castro *et al.*, 2005, 2010), there is still little empirical evidence that cultural adaptation positively affects outcomes when a program is delivered to different populations (Kumpfer *et al.*, 2002). Although the broader issue of cultural adaptation is beyond the scope of the present paper, we examine here one related aspect of fidelity: the effects of adherence when a program is delivered to groups that differ culturally from the original validation sample.

The present study

SFP is an evidence-based family intervention that has been shown effective in improving parenting skills in adult participants and decreasing the likelihood that youth participants will engage in substance use (Spoth *et al.*, 2001). The program is delivered in two-hour sessions, one night a week for seven weeks. For the first hour of each session, parents meet in one group and youth meet in a separate group. In the second hour, parents and youth meet together and engage in interactive exercises. The SFP curriculum contains instruction, discussion, and skills practice that emphasize warmth, communication, consequences, and setting clear limits for youth. Youth learn about managing stress, communicating with parents, resisting peer pressure, and other skills using similar processes (Spoth *et al.*, 2001).

The present study is an extension of a previous examination of global adherence to program content in a multi-site implementation of SFP (Cantu *et al.*, 2010). In the earlier study, there was high adherence to program content, and adherence was unrelated to short-term outcomes. Other studies of SFP fidelity have also used measures of global adherence to curriculum and found that fidelity was high (all implementations had greater than 80 percent adherence) (Fagan *et al.*, 2008; Spoth *et al.*, 2002, 2007) and unrelated to short-term outcomes (Spoth *et al.*, 2002, 2007). In those studies, researchers calculated a single adherence score representing the amount of material actually covered as a percentage of the total curriculum. The goal of the present study was to conduct a finer-grained analysis of the relation of specific program elements to short-term outcomes, by measuring adherence to program components – that is, we categorized program content into conceptually or procedurally distinct elements and calculated adherence within each category.

Our research questions were, first, whether adherence to specific program components was related to short-term outcomes, even though, in our original study (Cantu *et al.*, 2010), global adherence ratings were not related to outcome. Second, we wanted to know whether adherence to both scripted procedural elements (e.g. actively supervising a group activity) and scripted content (e.g. providing information about adolescent development) was related to program outcomes. Finally, given the emphasis in the literature on cultural adaptation, we

wanted to see whether effects of fidelity differed across two groups: the majority population in the region and in our participant sample (European American); and minority participants (primarily Latino/a and American Indian).

Method

Program observations

We conducted an observational study on programs delivered throughout Washington State during a six-month period. We asked providers we knew from trainings and previous implementations whether they would be willing to participate in an observational study of fidelity. Of 12 respondents who said they planned to implement a program, 11 agreed to be observed. We offered providers a \$250 incentive to be applied to program funding needs. Trained observers (a graduate student and the SFP state project evaluation coordinator) attended each program either singly or in pairs. When two observers were present, one observed the youth and family sessions and the other observed the parent and family sessions. Observers sat in the back of the room and completed implementation fidelity forms during the session.

Some facilitators felt that outsiders should not be present on the program's first night; thus, we did not observe night 1 but did observe the other six nights. In total, we observed 25 youth hours, 29 parent hours, and 34 family hours across the 11 program sites. Nested within the programs were 47 facilitators, 151 parent participants, and 144 youth participants. Six programs were hosted by schools, two by faith-based organizations, one by a substance abuse rehabilitation center, and two did not report host or sponsoring agencies.

Development of coding scheme

We followed an inductive coding process (Thomas, 2006) to categorize the entire scripted curriculum into categories. Two researchers coded the SFP material twice: on the first reading, they examined program content, deriving 13 categories (which we call "components") that described all content covered in the curriculum (Table I). Content components included information delivered to participants – for example, didactic instruction about matching consequences to behaviors (consequences); effective methods of communication (communication); and increasing family identity and connectedness (family unity). On the second reading, they examined scripted processes, independent of content, deriving nine categories that described all process components of the curriculum (Table II). Process components included actions of the facilitator designed to support delivery of curriculum – for example, providing instructions about how to complete an activity (instructions); setting up materials or arranging the room for an activity (setup); and eliciting discussion from participants (eliciting response). Having derived the component categories, they iteratively coded subsets of the curriculum and refined category definitions until they achieved adequate interrater reliability (80 percent agreement). They then coded the entire curriculum; final interrater reliability coefficients were $\kappa = 0.86$ for process and $\kappa = 0.72$ for content coding.

Sample

Of 151 adult participants, 59 percent were female, with an average age of 37.9 ($SD = 8.65$). In total, 68 percent of adult participants were white/European American, 11 percent Latino, 11 percent American Indian, 2 percent black/African American, 7 percent multiple ethnicity, and 3 percent other. Because of the small numbers in most race/ethnicity categories, we created a dummy variable to represent whether a person was of minority race/ethnicity (coded as 1) or white/European American (coded 0) rather than using specific race/ethnicity categories in analyses. Nine programs had at least one minority participant, and two programs had a majority of minority participants. The number of families in programs ranged from four to 26, with a mode of 8 and a mean of 12.4. In total, 68 percent of families attended five or more of the seven program nights.

Measures

Program level: scoring for component analysis—Each parent, youth, and family session comprises several modules, and each module comprises several activities. For example, most nights the parent session opens with a review of homework practice (module 1). In that module facilitators might ask parents if they had discussed behavioral consequences that week (activity 1); how they showed love to their child (activity 2); and whether they held a family meeting (activity 3). Each session contains somewhere between 30 and 98 separate activities. Using the fidelity form created by the program developers, observers rated each activity on a three-point scale as yes (completely covered; rated 2), some (attempted or partially covered; rated 1), and no (not covered; rated 0).

We computed each category's adherence score as a percentage of the total possible adherence to activities within that category across the entire program curriculum. For example, if a category contained ten activities total, delivered across seven nights, the maximum sum for adherence to that category would be 20 (each activity rated 2, for completely covered, multiplied by ten activities). If all ten of those activities were completely covered (scored 2), the fidelity score for that category would be 20/20, or 100 percent. If five of those activities were completely covered (scored 2) and the other five partially covered (scored 1), the fidelity score for that category would be 15/20, or 75 percent. In this way, we computed a fidelity score for each process and content category, resulting in 22 fidelity scores. This differs from previous SFP studies in two ways: first, those studies calculated a single adherence score for the total amount of material covered in the program (as opposed to adherence scores for individual components); and second, they scored activities as either 0 (not covered) or 1 (covered).

Participant level: parenting skills—We used a 13-item measure of intervention-targeted parenting attitudes and behaviors (ITPB) used in the original SFP research trial (Spoth *et al.*, 1995), designed to assess parenting skills targeted by SFP: family communication and enjoyment, emotion regulation, clarity and consistency of rules and discipline, and opportunities and rewards for youth involvement in prosocial behaviors. Items on the ITPB were rated on a five-point Likert-type scale with response options ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores representing more positive parenting behaviors. Participants completed the ITPB at both pretest and

posttest; positive change from pretest to posttest represents improvement in parenting skills assessed by the ITPB. Interitem reliability as assessed by Cronbach's α was 0.85 at both measurement occasions.

Data analysis

We conducted multilevel analyses to account for non-independence of participant observations within program (within-program intraclass correlation was 0.27). The outcome variable was participants' posttest score on the parent behavior scale, and process and content components were the predictors. Covariates included gender and minority/majority status (both of which had both been found to relate to the parenting outcome in Cantu *et al.*, 2010) as well as the ITPB pretest score. Cantu and colleagues found that program-level factors, including number of families, number of facilitators, and match of participant, and practitioner race/ethnicity, were not significant predictors of program outcome and thus were not included in the models. We also included interaction of gender and of minority/majority status with each process or content component to test for moderating effects. There were no significant interaction effects for gender X component, and we dropped these interaction terms from final models. We conducted a separate analysis for each component, and in order to conserve space we present coefficients, standard error, *t*-scores and probability from each full model only for the main variables of interest: each process or content component, minority status, and each moderator effect (Table III, complete tables for multilevel analyses of each component will be provided on request).

Results

Descriptive statistics

In Tables I and II, we present mean adherence values for each process and content component and their zero-order correlations with the ITPB change score (posttest minus pretest). Average adherence for content components was 81 percent (SD = 17 percent) and ranged from 66 to 91 percent. Average adherence for process components was 79 percent (SD = 14 percent) and ranged from 73 to 87 percent. Thus adherence was slightly higher, and slightly more variable, for content components than for process components. Correlations of content components with the change score ranged from -0.34 to 0.22 and averaged 0.16. Correlations of process components with the ITPB change score ranged from -0.28 to 0.31 and averaged 0.14.

Relation of content components to outcome

In multilevel analyses, nine of 13 main effects for content components and 11 interaction coefficients (minority status $\times$ component) were significantly related to the parenting outcome. Interaction effects of higher adherence to empathy, rules, involvement, and behavior categories were negatively related to outcome. The other significant interaction effects were positive.

We plotted each interaction term to look at the relation of high and low adherence to change on the ITPB for minority and majority participants separately (Aiken and West, 1991; Frazier *et al.*, 2004). In Figures 1 and 2, we present charts for high and low adherence

groups for all significant interactions of content components with minority/majority status. The relation of adherence to change on the ITPB was similar regardless of whether white/European American participants experienced high or low fidelity on any given component: their change on the ITPB was equivalent and positive. However, for minority participants, change scores for those who experienced low fidelity were different from those who experienced high fidelity. Furthermore, the differences were inconsistent across components— among minority participants, for some components higher fidelity was associated with more positive change (Figure 1), but for other component slower fidelity was associated with more positive change (Figure 2).

Relation of process components to outcome

For six of nine process components, both the main effect of component fidelity and the moderator effect of minority/majority status were significant. All significant main effects except supervision were positive: higher fidelity was associated with greater change. For supervision, higher fidelity was negatively associated with change. As with the content components, examination of interaction effects showed that majority participants in both high- and low-fidelity conditions showed equivalent levels of positive change (Figure 3). In contrast, for all process components except supervision, minority participants who experienced high fidelity showed greater positive change.

Summary

For majority participants, adherence to program components appeared to be unrelated to program outcome: both those in programs with low adherence to a component and those in programs with high adherence to a component showed equivalent and positive levels of change from pretest to posttest, and this was true for both content and process components.

For minority participants, however, effects of adherence on outcome were more variable. On all process components except one, high fidelity for minority participants was associated with markedly greater change than those for any other grouping (high- and low-fidelity conditions for majority participants and low-fidelity condition for minority participants). On about half the content components, higher fidelity was also associated with greater change, but on the others, higher fidelity was associated with lower change.

Discussion

The overarching goal of the present study was to determine whether assessing adherence to individual program components yielded more information than assessing adherence to the program curriculum globally. More specifically, we wanted to know whether adherence to at least some components was related to change on the parenting outcome, even though global adherence was not (Cantu *et al.*, 2010). We found that after controlling for program effects, and with the inclusion of covariates and moderator effects in our models, adherence was related to participant outcomes for most process and content components. Adherence to both content and process components appeared equally important. Some associations of adherence to outcome were positive and some negative, so at least for SFP, global assessment of adherence would be likely to obscure the relation of adherence to fidelity.

Furthermore, the association of adherence with outcome was consistently moderated by majority/minority status. Below we discuss this finding in greater detail.

The moderating effect of minority status

Different outcomes across groups could be due to variations in implementation, in how participants responded to the program, or both. SFP 10–14 was originally developed in rural Iowa and tested with a population that was 99 percent European American; the program rests on cultural norms and assumptions of the audience for which the program was intended and on which it was tested. In our community-based study, only 75 percent of participants identified themselves as European American. Thus, we consider it likely that most of the moderator effects we observed are related to mismatch of some parts of the program content and process with participants' culture. However, our observations about these moderator effects are speculative.

Moderation effects of adherence to program process components

For program processes, the positive effects of adherence to all but one component were higher for minority participants, and effects of adherence to two processes were especially strong: instruction (providing clear instructions for an activity and describing its purpose) and elaboration (illustrating or demonstrating a concept). It is possible that activities and concepts that were culturally familiar and straightforward to the validation sample in the clinical trial are less familiar to minority participants; in this case, close adherence to processes involving explanation, illustration, and organization would likely be more beneficial.

In contrast, higher adherence to the supervision component, when a facilitator ensures that an activity proceeds as scripted, was negatively related to outcome for minority participants. It is possible that facilitators high in adherence to supervision were biased and therefore more intrusive or offered too much assistance to minority parents; or, because of historical patterns of majority dominance and discrimination, minority parents might be inclined to interpret supervision of their activities as intrusive or patronizing. Finally, it is possible that higher adherence to supervision of program activities occurs when content is less familiar or intuitive to participants – in other words, higher supervision might simply have been a proxy for poor cultural match of the material.

Moderation effects of adherence to program content components

The moderating effects of adherence to program content components are less consistent and more difficult to interpret. Let us first look at the component for which low fidelity was most negatively associated with change on the ITPB for minority parents, consequences. Other studies of the wider community-based SFP in Washington (Griffin, 2010; Stocker *et al.*, 2012) have shown that when asked which SFP tools are most useful to them, Latino and American Indian parents are less likely than European American parents to mention the application of consequences or, especially, of positive reinforcement in response to youths' behaviors. Thus, lower adherence to content related to applying consequences (perhaps because facilitators perceive a cultural mismatch of this content with their audience), and

less home review and practice of it, would result in a lower likelihood of behavior change on the ITPB, which directly assesses these parenting behaviors.

In contrast, for some content categories (e.g. empathy and involvement), lower fidelity was associated with positive change on the outcome scale for minority participants. The involvement component includes content related to including children in family decisions and in discussions about the establishment of rules or behavioral consequences. The empathy component includes content designed to increase parents' understanding of their children's perspectives and experiences. Participants from cultures that strongly emphasize respect for parents and elders may have a negative response to the concept of involving their adolescents in rule making. Similarly, participants from cultures who value collective over individual identity may have less interest in taking the perspective of their adolescents. Perhaps facilitators who are sensitive to discrepancies between cultural values and program content are also more likely to emphasize other, more culturally appropriate, material. In this case, the effect of low adherence to these content components may simply be that it makes room for higher fidelity on content more acceptable to participants.

Strengths and limitations

The adherence effects moderated by what we have designated minority/majority status are consistently present and significant, and we believe they are important. However, we emphasize caution in their interpretation. First, most of those in the minority group are from two different cultures, American Indian and Latino/a, with a very few individuals of other races or ethnicities. Second, within those two broader groupings of American Indian and Latino/a, there are multiple distinct subcultures. The sample is too small to allow for examination of differences by the larger race/ethnic groupings, much less by subgroup, and thus any conclusions we draw about culturally related responses to program content are necessarily speculative and general. Furthermore, it is not possible in this study to disentangle the cultural effects of race/ethnicity from those related to socioeconomic status. In addition, the possible explanations we present for moderator effects are *post hoc* and speculative. Finally, there is evidence of differing response styles and measurement variance between Latino and non-Latino participants on the parenting outcome measure in a larger statewide evaluation of SFP (Hill, 2013), which could explain some of the difference in association of outcomes with adherence in the present study.

However, we do believe that the moderator effects we found reflect the fact that the original SFP audience was culturally different from many of the participants in our study, and these cultural differences have implications for both researchers and providers. We turn to those implications after discussion of the study's strengths and limitations.

A strength of the present study is its geographically and ethnically diverse sample; a limitation, as noted above, is the small sample size, which prohibits more fine-grained analyses of subgroups. The use of observational methodology in a community-based implementation is a strength; most community-driven programs either do not assess fidelity, or they use self-report of fidelity from facilitators. Observers are less subject to social desirability bias, and observer ratings have been found to be more strongly related to outcomes (Lillehoj *et al.*, 2004). On the other hand, conducting observations could lead to a

selection bias; it is possible that providers from the single program who chose not to be observed were more or less likely to deviate from the program as specified, or different in other ways that may influence program implementation and outcomes. In addition, the mere presence of the observer could have been stressful for facilitators, causing them to deliver the program differently or with greater adherence than they normally would. Furthermore, as noted above, internal standards may change as a result of the intervention, leading to recalibration and response shift from pretest to posttest, and cultural differences in response styles may confound the interpretation of self-report at each measurement point and of the meaning of change from pretest to posttest (Rosenman *et al.*, 2011).

The measure used for observing and recording adherence was a strength in that it has been used by multiple studies of SFP, and the components coded in this study could therefore be replicated with existing observational data from other sites. Our study was also limited, however, by its reliance on this measure, since adherence is only one aspect of program fidelity, a multidimensional construct which may also include adaptation (the degree to which program content was changed, added to, or eliminated); dosage (the amount of programming participants received); quality of program delivery (qualitative aspects of the program delivery); and participant responsiveness (the degree to which participants are engaged in program content and activities) (Blakely *et al.*, 1987; Dane and Schneider, 1998; Greenberg *et al.*, 2001; Dusenbury *et al.*, 2003; Hansen *et al.*, 1988; Johnson *et al.*, 2010). For example, observers awarded two points if a program element was delivered as specified, one point if a program element was delivered somewhat as specified, and no points if the element was not delivered. Scoring adherence in this manner does not account for the possibility of improvement through modifying program elements.

Blakely *et al.* (1987) and McGrew *et al.* (1994) demonstrated that some adaptations may lead to better outcomes, and results of the present study showed that lower adherence to some components was related to better outcomes for some minority participants. As noted above, our inferences about why this might be were speculative. Measures of adaptation, delivery quality, and participant responsiveness would have provided information useful in interpreting these moderator effects.

Finally, we conducted separate analyses for each of the 22 components, which increases the possibility of family-wise error. At the 0.05 probability level, we would expect to find significant associations simply by chance in one (5 percent) of the analyses; however, there were significant effects in 16 (73 percent) of the tests, more than might be expected by chance.

Implications

Effects of adherence may not be evident in global assessment of fidelity. Some components of implementation matter more than others, and the importance of specific components may depend on participant characteristics. In particular, fidelity to processes such as explaining new concepts, providing examples, clear instructions, and organized setup may be especially important when delivering a program to an audience that differs substantially from the clinical trial sample. For this reason, examination of adherence to specific program components, and the interaction of adherence with race, ethnicity, SES, and other

characteristics, will become increasingly important in program evaluation as evidence-based interventions move to large-scale dissemination in real-world settings (Fox *et al.*, 2004; Sanders *et al.*, 2002).

The adherence-outcome picture becomes increasingly complex if implementation quality, dosage, participant responsiveness, facilitator characteristics, and other contextual issues are considered. Available explanations of program theory are rarely able to account for so complex a picture – written materials that accompany a program curriculum typically summarize program activities, themes of program content, and expected results. However, all activities of a program are intended to serve some purpose, and it would be possible for developers to provide an explicit program theory – that is, to provide a categorization of each program's content and processes as part of the printed curriculum rather than someone creating an *ex post facto* coding scheme, as we did. Program developers could devise a specific, detailed, and explicit logic model or other representation or explanation of which specific program activities are expected to lead to changes in behavior, or how program processes support those changes (Chinman *et al.*, 2005). This type of complete and detailed program theory could significantly improve practitioner training, and it would provide a framework to guide practitioners' decisions about how and when it would be helpful or harmful to adapt program material, especially with regard to adaptations for different cultural groups. Detailed program theory would also be helpful to evaluators of community-based programs, who would be better able to determine which ingredients of a program work, for whom, and under what conditions (as well as which ingredients are not critical to program success). This information would then provide valuable guidance to program providers in real-world disseminations, enabling greater flexibility of implementation and improved outcomes.

Acknowledgments

The study was supported in part by two grants from the National Institute of Drug Abuse of the US National Institutes of Health (R21 DA025139-01A1 and R21 DA19758-01). The authors' thanks to Angel Cantu, on whose research the present study is based; to Chelsea Larse and Brianne Hood for research assistance; to the program facilitators and families who participated in the program evaluation; and to the Washington State Department of Social and Health Services for their collaboration.

References

- Aiken, LS.; West, SG. Multiple Regression: Testing and Interpreting Interactions. Thousand Oaks, CA: Sage Publications Inc; 1991.
- Blakely CH, Mayer JP, Gottschalk RG, Schmitt N, Davidson W, Roitman DB, Emshoff JG. The fidelity–adaptation debate: implications for the implementation of public sector social programs. *American Journal of Community Psychology*. 1987; 15(3):253–268.
- Cantu AM, Hill LG, Becker LG. Implementation quality of a family-focused preventive intervention in a community based dissemination. *Journal of Children's Services*. 2010; 5(4):18–30.
- Castro FG, Barrera M Jr, Martinez CR Jr. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. *Prevention Science*. 2005; 5(1):41–46. [PubMed: 15058911]
- Castro FG, Barrera M Jr, Steiker LKH. Issues and challenges in the design of culturally adapted evidence-based interventions. *Annual Review of Clinical Psychology*. 2010; 6:213–239.
- Chinman M, Hannah G, Wandersman A, Ebener P, Hunter SB, Imm P, Sheldon J. Developing a community science research agenda for building community capacity for effective preventive

- interventions. *American Journal of Community Psychology*. 2005; 35(3–4):143–157. [PubMed: 15909791]
- Dane A, Schneider B. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review*. 1998; 18(1):23–45. [PubMed: 9455622]
- Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*. 2008; 41(3–4):327–350. [PubMed: 18322790]
- Dusenbury L, Brannigan R, Falco M, Hansen W. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research*. 2003; 18(2): 237–256. [PubMed: 12729182]
- Fagan AA, Hanson K, Hawkins JD, Arthur MW. Bridging science to practice: achieving prevention program implementation fidelity in the community youth development study. *American Journal of Community Psychology*. 2008; 41(3–4):235–249. [PubMed: 18302016]
- Fox D, Gottfredson DC, Kumpfer KL, Beatty P. Challenges in disseminating model programs: a qualitative analysis of the strengthening Washington DC families program. *Clinical Child and Family Psychology Review*. 2004; 7(3):165–176. [PubMed: 15645706]
- Frazier PA, Tix AP, Barron KE. Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*. 2004; 51(1):115–134.
- Gottfredson DC, Kumpfer K, Fox D, Wilson D, Puryear V, Beatty P, Vilmenay M. The strengthening families Washington DC families project: a randomized effectiveness trial of family-based prevention. *Prevention Science*. 2006; 7(1):57–73. [PubMed: 16555144]
- Greenberg MT, Domitrovich C, Bumbarger B. Preventing mental disorder in school-aged children: current state of the field. *Prevention & Treatment*. 2001; 4(2):1–64.
- Griffin, A. unpublished McNair thesis. Washington, DC: Washington State University; 2010. The responses of Latino and Caucasian parents to a family strengthening program.
- Hansen WB, Graham JW, Wolkenstein BH, Lundy BZ, Pearson JL, Flay BR, Johnson CA. Differential impact of three alcohol prevention curricula on hypothesized mediating variables. *Journal of Drug Education*. 1988; 18(2):143–153. [PubMed: 3404365]
- Hill, LG. Apples and oranges: (non)equivalence of measurement across cultures and times. Washington, DC: Washington State University; 2013. unpublished manuscript
- Johnson KW, Ogilvie KA, Collins DA, Shamblen SR, Dirks LG, Ringwalt CL, Norland JJ. Studying implementation quality of a school-based prevention curriculum in frontier Alaska: application of video-recorded observations and expert panel judgment. *Prevention Science*. 2010; 11(2):275–286. [PubMed: 20358287]
- Kalafat J, Illback RJ, Sanders D. The relationship between implementation fidelity and educational outcomes in a school-based family support program: development of a model for evaluating multidimensional full-service programs. *Evaluation and Program Planning*. 2007; 30(2):136–148. [PubMed: 17689320]
- Kumpfer K, Alvarado R, Smith P, Bellamy N. Cultural sensitivity and adaption in family-based prevention interventions. *Prevention Science*. 2002; 3(3):241–246. [PubMed: 12387558]
- Lillehoj CJG, Griffin KW, Spoth R. Program provider and observer ratings of school-based preventive intervention implementation: agreement and relation to youth outcomes. *Health Education and Behavior*. 2004; 31(2):242–257. [PubMed: 15090124]
- McDonnell MG, Rodgers ML, Short RA, Norell D, Pinter L, Dych DG. Clinician integrity in multiple family groups: psychometric properties and relationship with schizophrenia client and caregiver outcomes. *Cognitive Therapy and Research*. 2007; 31(6):785–803.
- McGrew JH, Bond GR, Dietzen L, Salyers M. Measuring the fidelity of implementation of a mental health program model. *Journal of Consulting and Clinical Psychology*. 1994; 62(4):670–678. [PubMed: 7962870]
- Orrell-Valente JK, Pinderhughes EE, Valente E, Laird RD. If it's offered, will they come? Influences on parents' participation in a community-based conduct problems prevention program. *American Journal of Community Psychology*. 1999; 27(6):753–783. [PubMed: 10723534]

- Orwin RG. Assessing program fidelity in substance abuse health services research. *Addiction*. 2000; 95(11s3):309–327.
- Rosenman R, Tennekoon V, Hill LG. Measuring bias in self-reported data. *International Journal of Behavioural and Healthcare Research*. 2011; 4(4):320–332.
- Sanders MR, Turner KMT, Markie-Dadds C. The development and dissemination of the triple p-positive parenting program: a multilevel, evidence-based system of parenting and family support. *Prevention Science*. 2002; 3(3):173–189. [PubMed: 12387553]
- Spoth R, Redmond C, Shin C. Randomized trial of brief family interventions for general populations: adolescent substance use outcomes four years following baseline. *Journal of Consulting and Clinical Psychology*. 2001; 69(4):627–642. [PubMed: 11550729]
- Spoth R, Guyll M, Trudeau L, Goldberg-Lillehoj C. Two studies of proximal outcomes and implementation quality of universal preventive interventions in a community-university collaboration context. *Journal of Community Psychology*. 2002; 30(5):499–518.
- Spoth R, Redmond C, Haggerty K, Ward T. A controlled parenting skills outcome study examining individual difference and attendance effects. *Journal of Marriage and the Family*. 1995; 57(2):449–464.
- Spoth R, Guyll M, Lillehoj CJ, Redmond C, Greenberg M. PROSPER study of evidence-based intervention implementation quality by community – university partnerships. *Journal of Community Psychology*. 2007; 35(8):981–999. [PubMed: 20376336]
- Stocker, L.; Crawford, J.; Shrestha, G.; Eaton, A.; Betz, DL.; Hill, LG. Cultural and Ethnic Differences in Experience of an Evidence-Based Prevention Program. Pullman, WA: Wiley Research Exposition, Washington State University; 2012.
- Substance Abuse and Mental Health Services Administration (SAMHSA). [accessed March 16, 2013] NREPP: SAMHSA's national registry of evidence-based programs and practices. 2012. available at: <http://nrepp.samhsa.gov/>
- Thomas DR. A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*. 2006; 27(2):237–246.
- Yancey AK, Ortega AN, Kumanyika SK. Effective recruitment and retention of minority research participants. *Annual Review of Public Health*. 2006; 27:1–28.

Figure 1.

Interaction of adherence to content categories and minority status

Note: This bar graph includes content categories for which high fidelity was associated with positive change from pretest to posttest on the intervention-targeted parent behavior scale for minority participants

Figure 2.

Interaction of adherence to content categories and minority status

Note: This bar graph includes content categories for which low fidelity was associated with positive change from pretest to posttest on the intervention-targeted parent behavior scale for minority participants

Figure 3.

Interaction of adherence to process categories and minority status

Note: This bar graph includes process categories for which low and high fidelity were associated with different outcomes for minority participants

Table I

Content codes: description, average adherence levels, and correlation with parent outcome

Label	Description	<i>M (SD)</i>	<i>r</i>
Consequences	The positive and negative results of behaviors	78% (10%)	0.22**
Communication	Techniques of effective communication or communication of love and warmth	85% (10%)	0.17**
Family unity	Promotion of the family or closeness of family	73% (18%)	0.13
Rules	Making, keeping, or characteristics of effective rules	85% (16%)	-0.34*
Behavior	Desired and undesired behaviors youth engage in or could engage in	84% (30%)	-0.19*
Emotional regulation	Importance of, or techniques for, controlling emotions in potentially volatile situations	84% (25%)	0.21*
Peer issues	Resisting negative influence of peers	83% (22%)	-0.12
Involvement	Opportunities for youth to be involved in making rules, consequences, or family decisions	91% (20%)	-0.32*
Goals	Discussions or activities regarding desires of youth for their futures	66% (18%)	0.01
Empathy	Activities or discussions that promote youth understanding of parents' experiences or perspectives or parent understanding of youths' experiences or perspective	84% (14%)	-0.30*
Application	Review or discussion of specific application of program material, or what is working well for parents in general	84% (15%)	-0.04
Group unity	Promote group cohesiveness or encouragement and enthusiasm	85% (15%)	0.03
Other	Items that do not fit into above categories	73% (13%)	-0.01

Notes: $n=96$. The parenting outcome is the change score (posttest-pretest) on the intervention-targeted parenting behavior scale (ITPB); higher scores indicate desired outcome. Adherence scores range from 0 to 1 and represent the percentage of total possible coverage of program content within the category. The total number of parent participants was 155, of whom 96 submitted both pretest and posttest; thus the effective sample size for correlations was 96.

* $p < 0.05$;

** $p < 0.10$

Table II

Process codes: description, average adherence levels, and correlation with parent outcome

Label	Description	<i>M (SD)</i>	<i>r</i>
Elaboration	Facilitator gives information to illustrate or demonstrate a principle	76% (15%)	0.19**
Overview/preview	Facilitator introduces, states the purpose of, or explains a topic the group will discuss or activity the group will engage in	83% (21%)	0.03
Homework	Facilitator encourages group to use skills or reflect on knowledge gained between sessions	79% (20%)	0.02
Set up	Facilitator physically organizes or reorganizes group or prepares group or materials for an activity	80% (12%)	0.31*
Eliciting response	Facilitator asks group members to verbally provide answers to questions, examples, or to discuss a principle or activity	79% (10%)	0.02
Instruction	Facilitator explains how to participate in an activity. Instructions must provide opportunity for group compliance	73% (13%)	0.21*
Supervision	Facilitator ensures that an activity or process occurs properly, or prevents an activity from proceeding improperly	77% (13%)	-0.28*
Ritual	Facilitator leads an activity that is repeated over multiple sessions with special meaning attached to it	87% (15%)	-0.09
Material utilization	Facilitator distributes or makes reference to written materials or visual aids	85% (9%)	0.09

Notes: The parenting outcome is the change score (posttest-pretest) on the intervention-targeted parenting behavior scale (ITPB); higher scores indicate desired outcome. Adherence scores range from 0 to 1 and represent the percentage of total possible coverage of program content within the category. The total number of parent participants was 155, of whom 96 submitted both pretest and posttest; thus the effective sample size for correlations was 96.

* $p < 0.05$;

** $p < 0.10$

Table III

Statistics from multilevel analyses of the effect of adherence to each program component with parenting outcome

Category	Content components				Process components				
	B	SE	t	p	B	SE	t	p	
Consequences	7.52	0.84	8.90	<0.001	Elaboration	3.99	1.03	3.86	0.004
Minority	-6.54	0.76	-8.63	0.013	Minority	-2.88	0.60	-4.79	0.009
Interaction term*	8.11	0.97	8.31	<0.001	Interaction term	4.07	0.95	0.431	<0.001
Communication	4.38	1.85	2.37	0.042	Overview/preview	-0.79	2.08	-0.38	0.711
Minority	-4.05	1.45	-2.80	0.049	Minority	0.41	1.69	0.24	0.819
Interaction term	4.46	1.74	2.56	0.012	Interaction term	-0.94	2.97	-0.45	0.652
Family unity	2.18	0.59	3.71	0.005	Homework	1.26	0.85	1.49	0.172
Minority	-1.85	0.38	-4.91	0.008	Minority	-1.34	0.58	-2.29	0.084
Interaction term	2.35	0.56	4.19	<0.001	Interaction term	1.40	0.82	1.72	0.089
Rules	-3.62	0.61	-5.92	0.004	Setup	4.13	0.78	5.28	<0.001
Minority	2.32	0.58	3.98	0.058	Minority	-3.29	0.53	-6.15	0.004
Interaction term	-3.37	0.69	-4.90	<0.001	Interaction term	3.98	0.71	5.65	<0.001
Behavior	-0.82	0.42	-1.97	0.080	Eliciting response	2.95	1.11	2.67	0.026
Minority	0.30	0.29	1.02	0.365	Minority	-2.89	0.85	-3.40	0.027
Interaction term	-0.91	0.38	-2.40	0.019	Interaction term	3.33	1.10	3.02	0.003
Emotion regulation	1.91	0.48	4.02	0.004	Instruction	4.39	0.88	5.01	0.001
Minority	-1.77	0.37	-4.84	0.008	Minority	-3.20	0.53	-6.01	0.004
Interaction term	1.89	0.46	4.10	0.001	Interaction term	4.44	0.81	5.51	<0.001
Peer issues	3.51	2.38	1.47	0.191	Supervision	-2.93	0.76	-3.84	0.004
Minority	-3.80	2.36	-1.61	0.248	Minority	1.63	0.61	2.67	0.056
Interaction term	3.78	2.39	1.59	0.119	Interaction term	-2.43	0.75	-3.24	0.002
Involvement	12.25	0.99	-2.27	0.151	Ritual	0.95	0.78	1.21	0.257
Minority	2.17	0.68	3.20	0.085	Minority	-1.13	0.66	-1.71	0.162
Interaction term	-3.07	0.75	-4.09	<0.001	Interaction term	0.89	0.74	1.20	0.234
Goals	2.22	0.74	2.99	0.017	Material use	4.34	1.04	4.19	0.002
Minority	-1.98	0.44	-4.51	0.020	Minority	-4.65	0.79	-5.88	0.004
Interaction term	2.70	0.73	3.68	<0.001	Interaction term	5.37	0.98	5.50	<0.001

Category	Content components				Process components			
	B	SE	t	p	B	SE	t	p
Empathy	-3.74	1.02	-3.68	0.005				
Minority	3.28	0.82	4.01	0.016				
<i>Interaction term</i>	-4.00	0.90	-4.46	<0.001				
Application	1.77	0.66	2.69	0.025				
Minority	-2.16	0.55	-3.91	0.017				
<i>Interaction term</i>	2.23	0.66	3.36	0.001				
Group unity	1.29	0.71	1.81	0.104				
Minority	-1.07	0.60	-1.77	0.151				
<i>Interaction term</i>	0.82	0.68	1.20	0.232				
Other	3.52	1.29	2.72	0.042				
Minority	-2.84	1.01	-2.82	0.106				
<i>Interaction term</i>	3.63	1.35	2.69	0.010				

Notes: $n=96$. "Interaction term" indicates the interaction of minority race/ethnicity status with program component adherence. Minority status is coded as 1 and non-minority status as 0. Coefficients of significant effects are bolded. Interactions of gender with adherence were also examined but were non-significant and not included in the models. Each model also included gender and the pretest value on the intervention targeted parenting behavior scale (ITPB); we do not include these in the table because of space limitations, but full model statistics are available from the first author