

---

# Fidelity Decision Making in Social and Behavioral Research: Alternative Measures of Dose and Other Considerations

*Tiffany Washington, Sheryl Zimmerman, John Cagle, David Reed, Lauren Cohen, Anna Song Beeber, and Lisa P. Gwyther*

---

This article advances the discussion of treatment fidelity in social and behavioral intervention research by analyzing fidelity in an intervention study conducted within participating long-term care settings of the Collaborative Studies of Long-Term Care. The authors used the Behavior Change Consortium's (BCC) best practices for enhancing treatment fidelity recommendations in the areas of study design, provider training, treatment delivery, treatment receipt, and treatment enactment to evaluate fidelity-related decisions. Modifications to the original fidelity strategies were necessary in all areas. The authors revised their dose score and compared it with two constructed alternative measures of fidelity. Testing alternative measures and selecting the best measure post hoc allowed them to observe chance differences in relationship to outcomes. When the end result is to translate behavioral interventions into real practice settings, it is clear that some degree of flexibility is needed to ensure optimal delivery. Based on the relationship of program elements to the outcomes, a multicomponent intervention dose measure was more appropriate than one related to individual elements alone. By assessing the extent to which their strategies aligned with the BCC recommendations, the authors offer an opportunity for social work researchers to learn from their challenges and decision-making process to maximize fidelity.

KEY WORDS: *dose; intervention research; long-term care; treatment fidelity*

A significant problem with social and behavioral intervention research is that *treatment fidelity*, the extent to which an intervention is delivered as intended (Glasgow, Vogt, & Boles, 1999), is often overlooked. Unfortunately, careful observation of fidelity continues to be a methodological and pragmatic challenge for numerous reasons (Fraser, 2004; Tucker & Blythe, 2008). Intervention researchers have noted that observing, recording, and evaluating intervention encounters is time consuming and expensive (Bell et al., 2004; Planas, 2008). In addition, many social and behavioral interventions are multifaceted, making it difficult to assess the quantity, quality, and timing of the intervention as a whole. Furthermore, when interventions consist of multiple components, it may not be clear which components are important to measure, and if they are, how to operationalize them. These and other reasons may explain why most studies (39% to 55%) do not use any fidelity safeguards (Borelli et al., 2005; Moncher & Prinz, 1991; Naleppa & Cagle, 2010; Tucker & Blythe,

2008). This oversight is concerning, because if treatment fidelity is not closely monitored, the results of the intervention may be uninterpretable (Dane & Schneider, 1998; Gearing et al., 2011). Such a lack of clarity may impede research translation when the ultimate goal of the intervention is to promote adoption.

To inform future social work intervention research, this article describes the decision-making process involved when implementing and evaluating a multisite program to increase family involvement and enhance staff communication in 12 long-term care (LTC) settings. LTC settings encounter similar implementation challenges as found in other social services settings, such as staffing issues (for example, high turnover), client or resident attrition, and frequent regulatory changes (Buckwalter et al., 2009). Furthermore, regardless of setting, outside interventionists must also consider the existing resources and space to conduct the program (Fletcher et al., 2010), and how intervention delivery is affected by limited staff time.

The Behavior Change Consortium's (BCC) five-part treatment fidelity framework is useful to conceptualize fidelity-related decisions, as it clearly identifies the important components of study design, provider training, treatment delivery, treatment receipt, and treatment enactment (Bellg et al., 2004; Borrelli et al., 2005). In adhering to this framework, our decision-making process was guided by two assumptions. First, that some flexibility is necessary in social and behavioral outcomes research because modifications to the design and implementation may be indicated when working in challenging environments and with diverse populations. Second, although dosing may be easily conceptualized and measured in pharmaceutical trials (Farmer, 1999; Fraser, Galinsky, Richman, & Day, 2009; Friedman, Furberg, & DeMets, 2010), social and behavioral intervention research requires some creativity and innovation to conceive of and measure dose, while concomitantly ensuring the flexibility needed to adapt to individual and site differences (Fraser et al., 2009).

## DESIGN AND METHOD

### Program Description

Families Matter in Long-Term Care is an educational intervention designed to improve family involvement and promote better resident, family, and staff outcomes, and was adapted from two existing caregiver interventions, Partners in Caregiving (Pillemer, Hegeman, Albright, & Henderson, 1998; Pillemer et al., 2003) and the Family Partnership Program (Murphy et al., 2000). It was implemented in six nursing homes and 18 residential care/assisted-living settings. Baseline and outcome data were collected by research assistants from family members and staff using measures including the Interpersonal Conflict (Pillemer & Moore, 1989) and the Guilt subscales of the Family Perception of Caregiving Role (FPCR) instrument (Maas et al., 2004), the 22-item version of the Zarit Burden Interview (ZBI) (Zarit, Reever, & Bach-Peterson, 1980), and the Burden Subscale of the Lawton Caregiving Appraisal measure (Lawton, Kleban, Moss, Rovine, & Glicksman, 1989).

### Program Components and Fidelity Strategies

A three-hour educational in-person workshop was presented to families and staff separately by a trained

interventionist. The workshop included both didactic and interactive activities to increase knowledge of effective communication and partnership techniques and to begin to identify meaningful family roles that staff could support so as to promote resident quality of life. Included in the workshop were activities such as scaling (that is, assessing on a scale from 1 to 10 their relationship with the staff and their family member's quality of life), and "I" statements (that is, a communication activity that teaches individuals how to express their feelings in a conflict situation without blaming other parties involved).

Within one week after the workshops, families were to meet individually with trained interventionists and LTC staff to create a role for themselves through an individualized service plan. Families were prompted to develop plans tailored to the unique needs and preferences of the resident (so that, for example, a resident who previously enjoyed tending a garden, but could no longer do so, might be cognitively and physically supported by family and staff to tend to plants in the setting during family visits). To enhance resident quality of life and promote ongoing, meaningful interactions with the resident, service plans addressed one or more of the following "miracle outcome" domains (that is, an area in which they desired to see their family member's quality of life improved): getting around (for example, walking), doing things (for example, gardening), eating well (for example, hosting a dinner party), looking good (for example, getting a manicure), and helping the community (for example, decorating a common area for a holiday). For families who could not attend the workshop, staff were to carry out service plan creation independently with the support (if desired) of a trained, master's-level social worker with a background in LTC. Each interventionist was assigned to the same setting for the duration of the study, to serve as a liaison as needed.

Several fidelity strategies were conceived at the design phase. First, a participant would receive a full dose of the intervention by attending the entire workshop and implementing a service plan. To encourage workshop attendance, letters were mailed to families and an announcement about the upcoming workshop was posted in the community newsletter. To track attendance, participants were asked to sign in. Also, all participants received a certificate of completion, and staff members who attended the workshop received continuing

education credits. During the workshop, participants practiced creating meaningful service plans. A supply list was provided to families to aid in the development of the service plans. These supplies were made available to families to ensure that they possessed the materials required to successfully perform the activities (for example, watering pots, pedometers, art supplies).

To track adherence to the service plans, families were to have ongoing contact with the interventionist by way of follow-up telephone calls at one month, three months, and five months after service plan development, and postcard reminders at months two and four. During the calls, participants would be asked whether a service plan was created; if not, why not; and if so, to what extent it was being followed as planned.

### **Analytic Strategies to Measure Fidelity**

Our measure of exposure to the core intervention components (that is, the dose score) was developed to account for variations among participants' workshop attendance, service plan creation, and service plan adherence (that is, of 231 family member participants, 72% attended a workshop session, 70% developed a service plan, and 65% did both; only 37% adhered to their service plan). The measure was determined by the extent to which each component of the intended intervention was completed for individual participants. Dose was scaled on a range from 0 to 1; workshop attendance received a score of 0 (did not attend) or 1 (did attend); service plan creation was scored as either 0 (did not create) or 1 (created); and the extent of adherence to the service plan assessed via the monthly telephone calls was scored 0 (not at all), 0.5 (somewhat), or 1 (completely), with these scores averaged over the occasions on which adherence was reported, resulting in a score covering the full range from 0 to 1. Having the measure's value range from 0 to 1 results in the desirable property that the parameter for the measure when it is used in a regression model expresses the effect on the outcome measure of a full dose of the intervention, or complete fidelity. Not attending a workshop, no service plan, and no adherence yielded a score of 0 on these components. If no data were available regarding adherence, but the individual had created a service plan (and thus had the opportunity to be adherent), a score equal to the mean score for participants with the same attrition status was assigned. The

scores on the three components were then averaged to produce one single score per participant. The mean dose among all participants was .60. This was the measure used in analyses to evaluate the effect of the intervention (Zimmerman et al., 2013).

For comparison purposes, two conceptually reasonable alternative measures of fidelity, which also ranged from 0 to 1, were constructed and compared by varying the elements and the weighting of the program attributes. The first of these added a fourth attribute, service plan reinforcement (that is, the number of contacts made to inquire about adherence), to the other three, weighting all of them equally (because the act of contacting participants about the extent of their adherence contributes in itself to the intervention). The second alternative used the same program attributes as the basic strategy but changed the weighting for service plan adherence, giving it a double weighting to equal the workshop and service plan combined (because adherence to the service plan was theoretically considered to be the most important component of the intervention). The alternative composite measures, and the individual attributes of the intervention, were compared with the original composite measure on the basis of their basic descriptive statistics and on the effect on the outcome measures.

### **RESULTS**

To promote fidelity, several key modifications were made to the research protocol. First, because not all families or staff were available when the workshops were provided, an abbreviated, one-on-one version of the workshop, offered in-person or by telephone, was created. For families who missed the full three-hour workshop, the abbreviated version was offered at the beginning of the service plan meeting. Second, because the staff did not have sufficient time to meet with families, the interventionists, rather than LTC staff, assumed the full role of implementation. Consequently, service plans were not always created in the presence of a staff member as originally intended. Third, follow-up telephone calls and postcard reminders were reduced to one call at month three, and to one mailing at month one, respectively, to reduce participant burden. Details of these and other modifications to the protocol are provided in Table 1.

In terms of the measurement of dose, we examined the relationship of the four individual elements to outcomes and compared the effects of the

**Table 1: Original Fidelity Strategies and Subsequent Modifications**

BCC Fidelity Category	Original Fidelity Strategy	Modification
<b>Study Design:</b> Plan for implementation setbacks; establish a fixed dose	<ul style="list-style-type: none"> <li>Participants to attend the full version of the workshop</li> <li>Workshop and service plan creation conducted in person</li> <li>Fixed dose based on workshop attendance and service plan creation</li> </ul>	<ul style="list-style-type: none"> <li>A shortened version of the workshop was created to accommodate schedules</li> <li>The shortened version of the workshop and service plan creation may be conducted by telephone to accommodate schedules</li> <li>Varied dose based on workshop attendance, service plan creation, the extent of adherence to the service plan, and service plan reinforcement</li> </ul>
<b>Provider Training:</b> Assess the training of treatment providers	<ul style="list-style-type: none"> <li>Implementation by LTC staff</li> <li>Create individualized service plans with the input of a staff member</li> </ul>	<ul style="list-style-type: none"> <li>Interventionist assumed the full role of implementation</li> <li>Individualized service plans may be created with interventionist in the absence of a staff person</li> </ul>
<b>Treatment Delivery:</b> Deliver the intervention as intended	<ul style="list-style-type: none"> <li>Staff to undergo rigid training</li> <li>Observe staff delivery</li> </ul>	<ul style="list-style-type: none"> <li>Interventionists underwent rigid training due to implementation change</li> <li>Observe interventionists' delivery</li> </ul>
<b>Treatment Receipt:</b> Establish a method to verify participants understood and could perform the behavioral skills prescribed in the intervention	<ul style="list-style-type: none"> <li>Direct observation of behavioral skills</li> </ul>	<ul style="list-style-type: none"> <li>Follow-up telephone calls to address barriers to treatment receipt</li> </ul>
<b>Treatment Enactment:</b> Establish a method to monitor and improve the ability of patients to perform treatment-related behavioral skills in real-life settings	<ul style="list-style-type: none"> <li>Follow-up telephone calls to assess adherence to service plans at months 1, 3, and 5</li> <li>Postcard reminders mailed at months 2 and 4 to enforce adherence</li> </ul>	<ul style="list-style-type: none"> <li>Follow-up telephone calls reduced to months 1 and 3</li> <li>Postcards mailed at month 1</li> </ul>

Note: BCC = Behavior Change Consortium.

original composite measure to the effects of alternative composite measures of dose (see Table 2). Among the individual dose elements, workshop attendance related to one of four outcomes (that is, Lawton Burden subscale); service plan creation and service plan reinforcement (not an original item) related to two of four outcomes (that is, ZBI and Lawton Burden subscale for service plan creation, and Family Perceptions of Guilt and Lawton Burden subscale for service plan reinforcement); and service plan adherence related to none. The *p* values ranged from .02 (workshop attendance to Lawton Burden subscale; effect size  $-2.2$ ) to .048 (service plan reinforcement to the Family Perceptions of Guilt and Lawton Burden subscale; effect sizes 0.28 and  $-2.1$ , respectively). Note that effect sizes cannot be compared across measures because they are not standardized.

The composite dose measures all had means toward the middle of the 0 to 1 range (.5 to .6). The distributions departed significantly from normality with negative kurtosis ( $-1.0$  to  $-1.1$ ,  $p < .001$  for all) and negative skew ( $-.6$  to  $-.8$ ,  $p < .001$  for all), reflecting an excess of 0 values and some

flatness to the peak of the distribution. All the composite measures yielded significant effects on all outcomes (*p* values ranged from .013 to .047).

Intercorrelations among the composite dose measures were very high (.97 to .99), with the full width of the 95% CI being  $\leq .015$  for all three intercorrelations. For example, the correlation between the four-element measure and the measure with adherence double weighted was .967 (95% CI = .960 to .974).

These high intercorrelations are consistent with their relationship to all of the outcome measures being similar to that of the original measure, with generally similar *p* values and estimates of effect. For example, the effect of having a score of 1 (full fidelity) on the four-element version of the fidelity measure (alternate 1) compared with having a score of 0 (no fidelity) was a relative decrease of 3.0 points on the ZBI ( $p = .037$ ), compared to a decrease of 2.5 points ( $p = .042$ ) for the same amount of variation on the original measure (standardized effects equaled 0.25 and 0.21, respectively).

All of the individual measures, which each express only a single aspect of fidelity, showed effects in the same direction as the estimates obtained using the

**Table 2: Descriptive Statistics and Effects on Outcomes for Alternative Measures of Fidelity (N = 231)**

Intervention Element and Composite Dose Measure	Comparison of <i>p</i> Values and Estimates of Effect on Outcome Measures <sup>b</sup>								
	<i>M</i> ( <i>SD</i> ) <sup>a</sup>	Family Perceptions of Caregiving Role (Conflict)		Family Perceptions of Caregiving Role (Guilt)		Zarit Burden Inventory		Lawton Burden	
		Effect (SE)	<i>p</i>	Effect (SE)	<i>p</i>	Effect (SE)	<i>p</i>	Effect (SE)	<i>p</i>
Individual elements									
Workshop attendance	0.72 (0.45)	0.19 (0.102)	.07	0.14 (0.115)	.24	-1.5 (1.58)	.35	-2.2 (0.94)	.020*
Service plan creation	0.70 (0.46)	0.10 (0.100)	.31	0.06 (0.113)	.57	-3.5 (1.54)	.024*	-1.9 (0.93)	.041*
Service plan adherence	0.37 (0.33)	0.14 (0.093)	.13	0.16 (0.111)	.15	-2.5 (1.6)	.11	-1.5 (0.83)	.07
Service plan reinforcement	0.26 (0.30)	0.17 (0.12)	.15	0.28 (0.139)	.048*	-3.5 (1.94)	.07	-2.1 (1.03)	.048*
Constructed composites									
Original dose measure—three elements <sup>c</sup>	0.60 (0.36)	0.17 (0.075)	.025*	0.20 (0.090)	.027*	-2.5 (1.24)	.042*	-1.6 (0.66)	.014*
Alternate 1 dose measure—four elements <sup>d</sup>	0.51 (0.32)	0.19 (0.86)	.028*	0.23 (0.102)	.024*	-3.0 (1.42)	.037*	-1.9 (0.76)	.013*
Alternate 2 dose with service plan adherence double weighted <sup>e</sup>	0.54 (0.34)	0.17 (0.080)	.033*	0.20 (0.096)	.038*	-2.7 (1.33)	.047*	-1.7 (0.71)	.018*

<sup>a</sup>Theoretical range is standardized at 0 to 1 for all measures, with a value of 1 equivalent to the full fidelity according to that measure. Actual range is 0 to 1 for all measures, except for Alternate 1, which had an actual range of 0 to 0.96.

<sup>b</sup>All effects are unstandardized, that is, in the original metric of the scale. Lower scores are better.

<sup>c</sup>Original three elements: workshop attendance, service plan creation, and service plan adherence.

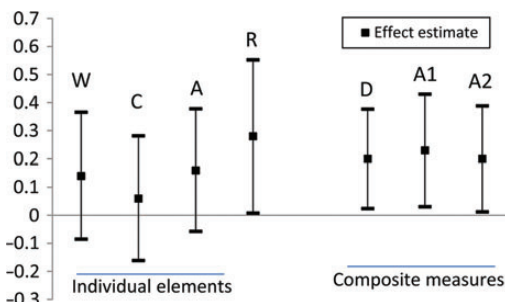
<sup>d</sup>First alternative measure adds a fourth element to the original three dose measures.

<sup>e</sup>Second alternative measure contains the same elements in the first alternative, but service plan adherence is double weighted.

\**p* < .05.

composite dose measures. Other features evident in the comparison among measures are illustrated

**Figure 1: Estimates of the Effect on the FPCR Guilt Subscale of Alternative Measures of Fidelity (Point Estimates and 95% Confidence Intervals)**



Note: FPCR = Family Perception of Caregiving Role; W = workshop attendance, C = service plan creation, A = service plan adherence, R = service plan reinforcement, D = original dose measure, A1 = alternative composite measure 1, A2 = alternative composite measure 2.

in Figure 1, which presents comparisons of the estimated effect sizes and 95% confidence intervals (CIs) for the relationship of the various fidelity measures to the FPCR Guilt subscale outcome. Similar patterns could be shown with any of the other outcomes. First, it is readily apparent that the size of the estimates varies across alternate measures of fidelity, but the 95% CIs for the estimates overlap greatly (more than 50% in all cases), indicating that the estimates do not differ significantly from one another across all the alternative measures. Second, the variation in size among the estimates is greater among the individual measures than among the composite measures. In the case of the FPCR Guilt subscale, the variation among the estimates from the individual measures is over five times greater than the variation among the estimates from the composite measures. Third, the 95% CIs were narrower for all of the composite measures than for the individual measures. This greater precision of the composite estimates explains why all of the relationships for the composite measures are statistically significant, whereas only a minority

of the estimates for the effect of the individual measures are significant. The CIs for the individual measures extend further, and therefore include zero more often, even though in general the estimates from the individual measures are comparable to those from the composites. This finding suggests that the composite measures reflect more accurate (that is, complete and reliable) information about the dose of the multi-component intervention than the individual measures, as would be expected if each makes a meaningful, distinct contribution to the intervention.

## DISCUSSION

Throughout the development and implementation of Families Matter, we were tasked with making decisions about fidelity promotion, including the optimal measurement of dose. Finding the balance between fidelity and flexibility can be challenging. In our case, most, but not all, decisions aligned with BCC's treatment fidelity strategy recommendations. For each of these domains, the considerations, decisions, and results related to ensuring fidelity in the Families Matter protocol are described in the following sections.

### Study Design, Provider Training, and Treatment Delivery Fidelity

Treatment fidelity strategies related to the design of the study ensure that the intervention's main ingredients are considered as the study is designed. This is generally achieved by determining the fixed dose to be given to each participant. However, in social and behavioral research, achieving a fixed dose is often beyond the ability of the interventionist, so it is especially important to maximize and monitor dose to inform the translation of research to practice (Durlak & DuPre, 2008; Fletcher et al., 2010; Mihalic, 2002) and allow for a better estimation of effect size. Individual and provider-related factors may affect dosing amount (the quantity of the intervention delivered at one point in time), frequency (the number of times the intervention is delivered over time), and duration (the entire length of time the intervention is delivered as a whole) (Reed et al., 2007). For example, dosing can vary based on provider knowledge, participant need (for example, participants with greater need may require a larger intervention dose), and participant ability to enact the intended behavioral skills.

In a multicomponent intervention such as Families Matter, a determination must be made about

which components to measure in regard to fidelity, and whether all components of the intervention are weighted equally. For example, workshop attendance was considered important to fidelity, because in both the abbreviated and the full versions, participants received useful information about improving their relative's quality of life. Adherence to the service plan was assumed to be the most integral component of the intervention, given that the project intended to promote family involvement through the enactment of service plans. We initially conceived of the dose as fixed (that is, service plan creation and workshop attendance), but as the project evolved, we recognized that dose needed to vary based on the extent to which a service plan was carried out. This variability was allowed because it became clear that family availability and other factors resulted in differences in service plan adherence.

Interpretations about what constitutes core ingredients in social and behavioral interventions vary widely (Gearing et al., 2011). Even with a fidelity protocol, it is sometimes difficult to pinpoint the "active" ingredients of an intervention. For example, we assumed service plan adherence was the most important program element given that our effort was to increase family involvement; however, adherence did not do better than any of the other individual measures in establishing an intervention effect on selected outcome measures. Also, the combined strategy that gave double weighting to service plan adherence did not yield greater effect sizes than either the four-element dose measure with equal weighting or the original dose measure. Testing such alternative strategies in preliminary work may help inform which components of a multicomponent intervention are integral. Of course, running analyses repeatedly to test alternative dose measures and then selecting the one with the best results should be avoided, because it involves taking advantage of chance (Harrell, 2001). Consequently, decisions about the dose measure in response to the realities of intervention implementation should be made independently of the results of analyses and tested in preliminary work.

The BCC recommends that researchers plan for implementation setbacks at the design stage. Ultimately, setbacks did occur. For instance, some family members were unable to attend the full workshops at the times they were offered. The shortened version that was offered as an alternative maintained the essential ingredients of the interventions (for example,



scaling, “I” statements, and miracle outcome activities), but eliminated introductions, background information about the development of the project, and group activities. Eliminating these activities allowed us to offer the shortened version by telephone if family members were unable to meet in person. We did not consider these activities integral to the intervention, and so were willing to modify it accordingly.

We decided which essential ingredients were important to maintain in the shortened version based on the theoretical assumptions guiding the intervention. For example, families were asked a miracle outcome question, “Considering what is important to the resident, what is the miracle outcome?” In answering this question, families were able to identify how they could add to the resident’s quality of life to aid in the development of a service plan—a key objective of the intervention. The miracle outcome question is a technique of solution-focused therapy, a goal-setting therapeutic approach based on social constructionism theory (Berg & De Jong, 1996). Having a theoretical basis for an intervention may help researchers make decisions related to fidelity when flexibility is required.

We also encountered roadblocks related to provider training. We could not expect staff to be readily available during the times the workshops were offered (especially not *all* staff at once), nor could we expect them to be available at the exact moment a service plan was being created. Therefore, we assumed the primary role of intervention implementation. Although there was no tangible measure of skill acquisition, as recommended by the BCC, it was obvious early during implementation that staff could not carry out the intervention as originally planned. This example illustrates an instance wherein actual capacity was not what was originally envisioned, and so might be ascertained by pilot study or more in-depth assessment of readiness for change.

To minimize differences in delivery, the interventionists underwent rigid training using scripted presentations and standardized materials, and workshop sessions were observed to ensure uniformity. Also, to reduce differences within treatment, each interventionist was assigned to the same site for the duration of the study.

### **Treatment Receipt and Treatment Enactment Fidelity**

The effectiveness of the intervention cannot be determined without a method of verifying that

the participants understood and could perform the behavioral skills prescribed in the intervention. Making follow-up telephone calls (which we referred to as the fidelity interviews and were a measure of service plan reinforcement) was the method we used for this purpose. These calls not only allowed us to verify that the information was understood, but also provided an opportunity to address barriers to treatment receipt—a strategy recommended by the BCC. For example, when participants were asked, “What barriers exist to making your service plan as successful as you would like?” the data collectors were instructed to refer the participant to the social work consultant.

The gold standard for assessing treatment enactment is direct observation; however, there was no way to predict exactly when the plan enactments would occur and could be observed. During the follow-up telephone calls, participants were asked whether a service plan was created; if not, why not; and if so, to what extent it was being followed as planned. The frequency of calls was eventually reduced because, as it turned out, some plans were created to be carried out once, and others were ongoing. Ascertainment of enactment therefore required flexibility, because service plans were individualized for each participant.

### **Implications**

In any social or behavioral intervention, it is the hope that participants actually perform the treatment-related behaviors in real-world settings. When translating behavioral interventions into real practice settings, it is clear that some degree of flexibility is needed to ensure optimal delivery (Cohen et al., 2008), although the question is to *what* degree. How much flexibility is needed to overcome implementation barriers (for example, limited resources) while maintaining the active ingredients of the intervention that are necessary to draw accurate conclusions about its effectiveness? For example, this project was originally intended to be more of a partnership between family and staff, with staff members fully involved in the development and implementation of the service plans. However, during implementation, staff availability varied, so their role became less central to the intervention. With the goal of assisting families to find ways to be more meaningfully engaged in their relatives’ care, the service plans remained a key intervention component.

The uptake and sustainability of interventions in real-life settings is fraught with barriers that prevent new programs from being incorporated into usual practices (for example, agency policies, informal practices, organizational culture) (Fraser et al., 2009). To identify unforeseen challenges, we encourage researchers to consider the benefit of expert review and pilot testing when developing program materials. Fraser et al. (2009) offered an extended discussion of these benefits and suggested that researchers closely examine the organizational or contextual constraints affecting implementation. Identification of these factors may result in necessary revisions to a program's format to increase the likelihood of adopting it into daily practices.

Although treatment fidelity strategies should be outlined before intervention implementation, researchers may not be able to fully predict potential setbacks. For instance, some fidelity strategies were developed from the onset (for example, follow-up telephone calls), but even with these strategies in place, additional measures (for example, dosing score) were developed to accommodate unforeseen challenges and opportunities.

How much program exposure is necessary to determine intervention efficacy? Efficacy cannot be determined if feasibility cannot be established. In a feasibility study, dosing decisions are determined by the essential ingredients of the intervention. The essential ingredients are derived from a clear theoretical model, which is critical to guiding research (Taylor & Bagd, 2005). The delivery of the intervention's active ingredients in the context of cost, fit, and ease of implementation in a real-life setting will determine the dosing needs to achieve a desired effect (Hohmann & Shear, 2002).

### Limitations

A noteworthy limitation to our approach is that family member report was used to determine if a service plan was completed and adhered to, which was used, in turn, to calculate dose scores. In the absence of true adherence, it is nonetheless possible, for example, that a family member may have answered "somewhat adherent" if there was intent to enact the service plan, or if there were barriers to completing the service plan. It is also possible that some service plans were too ambitious, so their relationship to outcomes might have been influenced by special challenges. Furthermore, there could have been other important reasons why service

plans were not implemented (for example, change in family member's health status), which themselves could have related to outcomes. Understanding the reason for the relationship of measures of fidelity to outcomes is beyond this study, but would be an important contribution to inform future research.

### CONCLUSION

By reporting our fidelity strategies and assessing the extent to which our strategies aligned with the BCC recommendations, we offer an opportunity for social work researchers to learn from the challenges we faced and the decisions we made to maximize fidelity throughout the development and implementation of Families Matter. Our findings have important ramifications for replicating studies in clinical settings (Borrelli et al., 2005). Also, we hope to encourage researchers to establish creative ways to calculate dose scores based on implementation barriers that inevitably occur, and to help identify the operative components of the intervention that relate to outcomes. **SWR**

### REFERENCES

- Bellg, A., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D., Ory, M., et al. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychology, 23*, 441–451.
- Berg, I. K., & De Jong, P. (1996). Solution-building conversations: Co-constructing a sense of competence with clients. *Journal of Contemporary Human Services, 77*, 376–391.
- Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A. J., Czajkowski, S., Breger, R., et al. (2005). A new tool to assess treatment fidelity across 10 years of health behavior research. *Journal of Clinical Psychology, 73*, 852–860.
- Buckwalter, K., Grey, M., Bowers, G., McCarthy, A., Gross, D., Funk, M., & Beck, C. (2009). Intervention research in highly unstable environments. *Research in Nursing and Health, 32*, 110–121.
- Cohen, D., Crabtree, B., Etz, R., Balasubramanian, M., Donahue, K., Leviton, L., et al. (2008). Fidelity versus flexibility: Translating evidence-based research into practice. *American Journal of Preventive Medicine, 35* (Suppl. 1), S381–S389.
- Dane, A., & Schneider, B. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23–45.
- 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, 41*, 327–350.
- Farmer, K. (1999). Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. *Clinical Therapeutics, 21*, 1074–1090.
- Fletcher, S., Zimmerman, S., Preisser, J., Mitchell, M., Reed, D., Gould, E., et al. (2010). Implementation fidelity of standardized dementia care training program



- across multiple trainers and settings. *Alzheimer's Care Today*, 11, 51–60.
- Fraser, M. (2004). Intervention research in social work: Recent advances and continuing challenges. *Research on Social Work Practice*, 14, 210–222.
- Fraser, M., Galinsky, M., Richman, J., & Day, S. (2009). *Intervention research: Developing social programs*. New York: Oxford University Press.
- Friedman, L., Furberg, C., & DeMets, D. (2010). *Fundamentals of clinical trials*. New York: Springer.
- Gearing, R., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2011). Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation. *Clinical Psychology Review*, 31, 79–88.
- Glasgow, R., Vogt, T., & Boles, S. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health*, 89, 1322–1327.
- Harrell, F. E. (2001). *Regression modeling strategies*. New York: Springer-Verlag.
- Hohmann, A. A., & Shear, M. K. (2002). Community-based intervention research: Coping with the noise of real life in study design. *American Journal of Psychiatry*, 159, 201–207.
- Lawton, M. P., Kleban, M. H., Moss, M., Rovine, M., & Glicksman, A. (1989). Measuring caregiving appraisal. *Journal of Gerontology: Psychological Sciences*, 44, 61–71.
- Maas, M. L., Reed, D., Park, M., Specht, J. P., Schutte, D., Kelley, L. S., & Buckwalter, K. C. (2004). Outcomes of family involvement in care intervention for caregivers of individuals with dementia. *Nursing Research*, 53, 76–86.
- Mihalic, S. (2002). *The importance of implementation fidelity*. Boulder, CO: Center for the Study and Prevention of Violence.
- Moncher, F., & Prinz, R. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, 11, 247–266.
- Murphy, K. M., Morris, S., Kiely, D. K., Morris, J. N., Belleville-Taylor, P., & Gwyther, L. (2000). Family involvement in special care units. *Research and Practice in Alzheimer's Disease*, 4, 229–239.
- Naleppa, M., & Cagle, J. (2010). Treatment fidelity in social work intervention research: A review of published studies. *Research on Social Work Practice*, 20, 674–681.
- Pillemer, K., Hegeman, C., Albright, B., & Henderson, C. (1998). Building bridges between families and nursing home staff: The partners in caregiving program. *Gerontologist*, 38, 499–503.
- Pillemer, K., & Moore, D. W. (1989). Abuse of patients in nursing homes: Findings from a survey of staff. *Gerontologist*, 29, 314–320.
- Pillemer, K., Sutor, J. J., Henderson, C., Meador, R., Schultz, L., Robison, J., & Hegeman, C. (2003). A cooperative communication intervention for nursing home staff and family members of residents. *Gerontologist*, 43, 96–106.
- Planas, L. (2008). Intervention design, implementation, and evaluation. *American Journal of Health-System Pharmacy*, 65, 1854–1863.
- Reed, D., Titler, M., Dochterman, J., Shever, L., Kanak, M., & Picone, D. (2007). Measuring the dose of nursing intervention. *International Journal of Nursing Terminologies and Classifications*, 18, 121–130.
- Taylor, A. C., & Bagd, A. (2005). The lack of explicit theory in family research: A case analysis of the *Journal of Marriage and the Family* 1990–1999. In V. L. Bengtson, A. C. Acock, K. R. Allen, P. Dilworth-Anderson, & D. M. Klein (Eds.), *Sourcebook of family theory & research* (pp. 22–25). Thousand Oaks, CA: Sage Publications.
- Tucker, A. R., & Blythe, B. (2008). Attention to treatment fidelity in social work outcomes: A review of the literature from the 1990s [Research Note]. *Social Work Research*, 32, 185–190.
- Zarit, S. H., Reeve, K. E., & Bach-Peterson, J. (1980). Relatives of the impaired elderly, correlates of feelings of burden. *Gerontologist*, 20, 649–655.
- Zimmerman, S., Cohen, L. W., Reed, D., Gwyther, L. P., Washington, T., Cagle, J. G., et al. (2013). Families matter in long-term care: Results of a group randomized trial. *Seniors Housing & Care Journal*, 21(1), 3–20.

---

**Tiffany Washington, PhD, MSW**, is assistant professor, School of Social Work, University of Georgia, 310 East Campus, Athens, GA 30602-7016; e-mail: twashing@uga.edu.  
**Sheryl Zimmerman, PhD, ACSW**, is distinguished professor and associate dean for doctoral education, School of Social Work, University of North Carolina at Chapel Hill. **John Cagle, PhD**, is assistant professor, School of Social Work, University of Maryland, Baltimore. **David Reed, PhD**, is analyst, and **Lauren Cohen, MA**, is research associate and associate director, Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill. **Anna Song Beeber, PhD**, is assistant professor, School of Nursing, University of North Carolina at Chapel Hill. **Lisa P. Gwyther, MSW, LCSW**, is director of Family Support Program and associate professor, Duke Center for the Study of Aging and Human Development, Duke University, Durham, North Carolina.

Original manuscript received January 18, 2013  
 Final revision received May 20, 2013  
 Accepted June 3, 2013  
 Advance Access Publication July 16, 2014