

INVITED PAPER

The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials

Belinda Borrelli, PhD

Centers for Behavioral and Preventive Medicine, The Miriam Hospital; and Warren Alpert School of Medicine at Brown University, Providence, RI

Keywords

research design; reproducibility of results; reliability and validity; treatment fidelity; internal validity; external validity; confounders; contamination.

Correspondence

Dr. Belinda Borrelli, Centers for Behavioral and Preventive Medicine, One Hoppin Street, Providence, RI 02903. Tel.: 401-793-8040; Fax: 401-793-8078; e-mail:

belinda_borrelli@brown.edu. Belinda Borrelli is a Professor at the Centers for Behavioral and Preventive Medicine, The Miriam Hospital and Warren Alpert School of Medicine at Brown University.

doi: 10.1111/j.1752-7325.2011.00233.x

Abstract

Objectives: To discuss methods of preservation of treatment fidelity in health behavior change trials conducted in public health contexts.

Methods: The treatment fidelity framework provided by the National Institutes of Health's Behavioral Change Consortium includes five domains of treatment fidelity (Study Design, Training, Delivery, Receipt, and Enactment). A measure of treatment fidelity was previously developed and validated using these categories.

Results: Strategies for assessment, monitoring, and enhancing treatment fidelity within each of the five treatment fidelity domains are discussed. The previously created measure of treatment fidelity is updated to include additional items on selecting providers, additional confounders, theory testing, and multicultural considerations.

Conclusions: Implementation of a treatment fidelity plan may require extra staff time and costs. However, the economic and scientific costs of lack of attention to treatment fidelity are far greater than the costs of treatment fidelity implementation. Maintaining high levels of treatment fidelity with flexible adaptation according to setting, provider, and patient is the goal for public health trials.

Introduction

Treatment fidelity is the ongoing assessment, monitoring, and enhancement of the reliability and internal validity of a study (1). Treatment fidelity helps to increase scientific confidence that the changes in the dependent variable (outcome of interest) are due to manipulations of the independent variable (presumed to have an effect on the dependent variable). Treatment fidelity consists of two general components: a) treatment integrity, the degree to which a treatment is implemented as intended, and b) treatment differentiation, the degree to which two or more study arms differ along critical dimensions (2-5).

Conclusive statements about treatment effects cannot be made without attention to treatment fidelity. For example, without assessment of treatment fidelity, significant results may be a function of either an effective intervention or the influence of other unknown factors added into (or omitted from) the intervention. The danger of this is type 1 error (belief that a treatment effect is significant when it is not)

and the potential for dissemination of ineffective treatments. Similarly, if treatment fidelity is not measured and there are nonsignificant effects, it cannot be known whether these effects are due to an ineffective treatment or to the omission or addition of potentially active or inactive components. The danger of this is type 2 error (erroneous belief that a treatment effect is nonsignificant) and the potential for discarding effective treatments (2,6). Thus, treatment fidelity enhances both the internal validity (the treatment is delivered as intended) and external validity (the treatment can be replicated and applied in real-world settings).

Rejection of effective programs or acceptance of ineffective programs due to lack of treatment fidelity has untold costs, both financially and to science. If fidelity is not measured during treatment delivery, increased costs may be incurred when independent labs attempt to replicate the original results but are unable to do so because the components of the treatment as actually delivered are unknown. Further costs are incurred if ineffective treatments are disseminated into

standard practice. A scientific cost of inferior treatment fidelity is that investigators may unwittingly try to build their careers on results that have little empirical basis (i.e., positive findings could be due to variables other than those specified in the intervention). The current paper will discuss the assessment, monitoring, and enhancement of treatment fidelity in public health trials, with examples from oral health and other health behavior change studies.

Benefits of treatment fidelity

Treatment fidelity allows for the early detection of errors to prevent protocol deviations from becoming widespread and long lasting, which can potentially affect the study's ultimate conclusion. Monitoring treatment fidelity early in study implementation increases the fidelity of implementation (7). High levels of treatment fidelity improve treatment retention and reduce attrition (8). Treatment fidelity is particularly important for cross-site studies, to ensure that treatments are operationalized (defining what is, and what is not, part of the treatment), and, in the same way, across sites and reducing the possibility of site by treatment interactions (9,10).

Treatment fidelity facilitates theory testing (11,12). High levels of treatment fidelity are associated with changes in the mediating variables (mechanisms of change) hypothesized to be responsible for study outcomes (13,14). Interventions that adhere more closely to theory have stronger effects (15). Simply articulating a theory without monitoring fidelity to the theoretical components is associated with weakened treatment effects (16).

Treatment fidelity implementation should, itself, have treatment fidelity. If one treatment is implemented more purely than another, then treatment condition differences may be due to differences in fidelity, rather than to treatment content. For example, if treatment fidelity is only measured in the experimental group, it is difficult to determine whether or not the control group received an active treatment ingredient (treatment component hypothesized to be strongly associated with outcome) from the experimental condition, or received some other active intervention component. This could have the effect of reducing the effect size between the treatment and control groups, leading the researcher to incorrectly conclude that the experimental treatment is not effective when it was actually not given a fair test. Similarly, without monitoring fidelity in the control group, it cannot be determined whether an iatrogenic component was added that had the effect of reducing change in the control group, thus artificially enhancing the differences between the two groups.

Higher levels of treatment fidelity are associated with better treatment outcomes (17). High-fidelity programs outperform low-fidelity programs (6,12,18) and poor fidelity

attenuates outcomes (19). One study found that higher levels of treatment fidelity were associated with greater improvement in diabetic regimen adherence and greater improvement in metabolic control among adolescents with diabetes (14). Furthermore, using structural equation modeling, a completely mediated pathway was found between treatment fidelity and metabolic control, with regimen adherence mediating this relationship. Improved study outcomes due to treatment fidelity are likely the result of reduction of random and unintended variability, which increases power to detect effects.

Maximizing treatment fidelity

My colleagues and I of the National Institutes of Health's (NIH) Behavioral Change Consortium (BCC) developed a comprehensive treatment fidelity framework tailored to be relevant for health behavior change trials (1,11,20). These best practice recommendations put forth guidelines for treatment fidelity across five domains: Study Design, Provider Training, Treatment Delivery, Treatment Receipt, and Treatment Enactment. Guidelines and strategies for assessing, monitoring, and enhancing treatment fidelity within each of these domains are discussed below. Appendix I displays a checklist that can be used to assess the treatment fidelity of a study across each of these five domains.

Study design

Principles

Treatment fidelity practices related to study design ensure that a study adequately tests its hypotheses in relation to its underlying theoretical and clinical processes (11). This involves operationalizing the treatment in such a way that treatment components are reflective of, and mapped onto, the theory. The hypothesized active ingredients of the treatment (those that are hypothesized to affect outcome) are made explicit in the treatment protocol, and in the training and follow-up supervision of providers.

Assessment of fidelity to study design

Prior to study implementation, investigators, and optimally a protocol review group or panel of experts, should review their protocols or treatment manuals to ensure that the active ingredients of the intervention are fully operationalized. The degree to which the measures reflect the hypothesized theoretical constructs and mechanisms of action should also be assessed. Using a protocol review group to ensure that the study design is operationalized as hypothesized is particularly important if the intervention is to target a specific population (ethnic, underserved, etc.). In

that case, the protocol needs to be evaluated further for cultural relevancy, and optimally, members from the target community should be involved in the design and implementation of the study, in line with community-based participatory research (21,22).

Investigators should also conduct a critical inventory of their study design, asking what might challenge the hypothesized causal influence between the Dependent variable (DV) and Independent Variable (IV) (e.g., that changes in the IV cause changes in the DV). For example, is there a control for contact time between treatment conditions, and if not, how will the study's conclusions be affected? A priori specification of treatment dose should be delineated for each condition, including the length of each contact, the number of contacts, duration of contact over time (length of time of intervention period), and treatment content. While a fixed dose of treatment is preferable, a minimum and maximum amount of treatment dosage (range) can be given to providers in clinical settings to allow for some flexibility.

Setbacks in study implementation could also confound study results. For example, unanticipated provider dropout could lead to hurried attempts at training new providers, possibly resulting in performance differences between new and existing providers. It is recommended that studies have access to a larger pool of providers and train backup providers from

the outset. These treatment fidelity assessment criteria are included on the checklist in Appendix I.

Monitoring

Monitoring to assess adherence to the original study design should be conducted at the beginning of study implementation and over the course of the study in order to prevent drift from the protocol. A plan should be developed for how the monitoring will occur (frequency and process), how protocol deviations will be recorded, and how feedback will be given to providers. Although monitoring fidelity to intervention delivery is discussed in a later section, it should be discussed here that part of monitoring involves ensuring theoretical fidelity, that the theory is adequately reflected in intervention delivery during all phases of the trial. One of the most stringent ways to monitor theoretical fidelity would be to have outside raters listen to the intervention, guess the underlying theory, and rate the presence or absence of the specified theoretical components.

Treatment dose should also be monitored over time, both within and between groups. Providers fill out a brief "intervention checklist" after each participant contact, indicating the length of the visit and the components delivered. Audiotaping or videotaping the encounter is the most objective way to assess length of visit and fidelity to intervention content.

Table 1 Methods of Monitoring Treatment Fidelity: Pros and Cons

Method	Pros	Cons
Audiotaping	Enables objective evaluation of treatment content and dosage. Coders rate adherence to the protocol. Allows for specific feedback to providers during supervision. Enables providers in training to listen to previous visits. Ensures standardization within and between providers. Digital recorders are inexpensive and data can be stored on an external hard drive.	Slightly obtrusive, though when framed as "quality control for the best care possible" the vast majority of participants agree to audiotaping. Both the control and the intervention groups should be monitored, and taping may influence the participant in unknown ways.
Videotaping	Has the same advantages as audiotaping, although videotaping enables the evaluation of nonverbal behaviors in both provider and patient.	More obtrusive and costly; may increase demand characteristics.
Provider self-report (checklist)	Serves as a reminder to providers the active ingredients to be delivered. Cues providers to implement treatments with fidelity. Providers might be more likely to deliver treatment components if they know they have to check off a "no" if they don't deliver the component. Self-report data can be used as a supplement to direct methods of assessment, and both methods can be compared to each other. Affords immediate access to integrity data.	Takes more provider time than audiotaping. Potential for providers to rate themselves as more adherent than they really are. Low agreement between self-report and observational methods.
Participant self-report questionnaire	Enables assessment of whether participants received the required treatment components or contraindicated components. Assess nonspecific process issues (patient the provider, patient felt listened to versus rushed, patient felt understood versus uncomfortable, and patient felt respected versus criticized). Patient satisfaction with treatment and perceptions of treatment effectiveness can also be assessed.	Subject to memory bias and accuracy. Participants may not want to give bad ratings to providers. Participants may not have the knowledge or training to describe what happened at the visit at the level needed for analysis.

Table 2 Methods of Enhancing Treatment Fidelity: Study Design

1.	Explicitly identify and use a theoretical model as a basis for the intervention, and ensure that the intervention components and measures are reflective of underlying theory. Use a protocol review group.
2.	Pilot test the intervention and use feedback from participants and providers to refine adherence to the theoretical model and improve acceptability, feasibility, and potential effectiveness of the intervention.
3.	Determine a priori the number, length, and frequency of contacts, and develop a monitoring plan to maintain consistency in dose.
4.	Develop a plan for how adherence to the protocol will be monitored (audiotaping, videotaping). Monitor both intervention delivery and assessment administration (to ensure consistency of measurement).
5.	Develop a plan to record protocol deviations (dose, treatment content) across all conditions and method of providing timely feedback to providers.
6.	Develop a user-friendly scripted curriculum or treatment manual (print or via computer/handheld device) to ensure consistency of delivery and adherence to active ingredients of the treatment.
7.	Plan for implementation setbacks (e.g., attrition of treatment providers). Videotape the trainings to ensure consistency for future trainings.

The pros and cons of different methods of fidelity monitoring are in Table 1. The strategies for enhancing treatment fidelity in the design of studies are listed in Table 2.

Preserving fidelity in provider training

Principles

Treatment fidelity of provider training involves standardizing training between providers, ensuring that providers are trained to criterion, and monitoring and maintaining provider skills over time. Ensuring treatment fidelity during training is mutually exclusive from that of study design: despite a perfectly operationalized study and protocol that adheres to underlying theory, if providers are not adequately trained and monitored, nonsignificant results at the end of the study could be due to either poor training or to an ineffective intervention. Well-trained providers are less likely to deviate from the treatment and are more likely to show increased competency to deliver the intervention (4,23).

At the study outset, it is important to develop a comprehensive training plan that includes the specification of provider characteristics to look for when hiring, and a plan for how to train them to criterion and help them maintain skills over time. The training plan should be driven by the treatment protocol, emphasize the theoretical underpinnings of the intervention, articulate the necessary knowledge and skill requirements required for effective treatment delivery, and identify appropriate resources for training (20).

When hiring providers, there should be some assurance that there is a good fit between the provider and the population (e.g., matching on ethnicity or gender), as well as a good fit between the provider and the treatment. The treatment should be described in detail to prospective providers to assess whether or not they perceive it to be credible and consistent with their own values. For example, it would be detrimental to hire a provider for a study on reducing alcohol use if the provider believed in “abstinence only” treatments. It is

also important to determine that providers are willing to be randomized to either a treatment or a control group, and that if they are randomized to the latter, that they will not be compelled to provide extra components.

Deciding a priori on the level of credentials and years of experience required for providers will help prevent unintended variation in outcomes (24). Consistency across providers in these background characteristics helps to prevent provider by treatment interactions (24).

If the intervention is occurring within a larger entity, such as a community clinic, it is important to obtain “buy in” from the overall organizational structure. Provider perception of organizational support has been shown to be critical for motivating provider counseling (25). Factors to consider when hiring providers are displayed in Table 3.

Supervisors should also be chosen carefully, demonstrating both the knowledge and expertise in the content areas targeted by the treatment. Supervisors should be rated by a national expert in order to maintain their own skills (e.g., x% of all provider sessions rated by the supervisor are co-rated by expert supervisors) (10). Furthermore, supervisors should be able to demonstrate that they have the requisite skills to facilitate the process of supervision, for example, providing feedback in an appropriate manner so as to not elicit defensiveness in trainees.

The training should be standardized to ensure that all providers are given the same amount of training and that training is consistent within and between providers (Table 3). This increases the likelihood that the intervention will be delivered systematically across providers and decreases the likelihood that there will be provider by treatment interactions (that treatment delivery varies between providers). Standardization of training, however, does not preclude individualization. Training needs to take into account the different levels of education, experience, learning styles, and counseling styles of different providers. Providers should also be taught how to deal with different types of patients (e.g., resistance).

Role playing with standardized patients and scoring the interaction on both adherence to process and content is one

Table 3 Methods of Enhancing Treatment Fidelity: Training

-
- Hiring: Hire providers with similar credentials and experience. Ensure “buy in” to treatment, theory, and randomization. Consider matching providers to key characteristics of the population.
 - Standardize training: Use the same trainers over time, use certified trainers, train all providers together, use standardized training materials, use video or audiotapes of expert delivery, develop a manual of training procedures and videotape trainings in case of provider attrition and need for future trainings.
 - Accommodate learner differences: Design training for diverse learning styles, train providers to deal with different types of participants, consider more intensive training and follow-up for less experienced providers.
 - Assess skill acquisition: Use role plays with standardized patients followed by feedback to provider, score provider adherence to both intervention content and process using validated performance criteria, have a written exam pre- and post-training, develop criteria for initial certification.
 - Prevent skills drift: Booster sessions, patient exit interviews, periodic re-certification, audio or video record all encounters and code for treatment adherence, provide timely feedback, monitor patient dropout rates of each provider.
 - Enhance buy-in from providers: Foster provider self-efficacy and perception of organizational support. Explain the study design and rationale, the principles of research, and why it is important to prevent contamination and omission or addition of components not specified by the intervention.
-

strategy to impart skills. Although many providers are reticent to complete role-plays, they often feel more confident to deliver the treatment with study participants after reaching competency during role-plays. One study showed that nurses' self-efficacy to provide smoking cessation counseling significantly increased after training, and was maintained at a 6-month follow-up (25). Other strategies for training are listed in Table 3.

Training should aim to foster meta-competence, ensuring that providers not only understand the treatment components but also the rationale and theory behind them. This increases a providers' ability to work flexibly with different patients while maintaining adherence to the study and the underlying theory (26). For example, if the intervention is based on the health belief model, and the provider encounters a novel situation, the provider could reference the theory and ensure that his or her response is consistent with the theory. Providing interventionists with an explanation of the rationale for treatment fidelity is important as well. This may include a discussion of the importance of preventing treatment contamination, and why it's important not to add components (even though they think they might be useful) or delete components (even though they think they might be ineffective). Criteria to evaluate treatment fidelity of training plans are presented in Appendix I.

Assessment of training

Assessment of training involves ensuring that providers are trained to a well-defined, a priori performance criterion. Provider role-plays with standardized patients should be evaluated for both adherence to treatment components and adherence to process (e.g., interactional style). A list of the theory-based active treatment ingredients should be developed, as well as a method to determine the degree to which they were implemented as intended (e.g., rating the use of

each treatment component on a 5-point Likert scale or simply rating the presence or absence of each treatment component). While there is no gold-standard cutoff for determining minimum trainee competency levels (e.g. adherence to ≥ 95 percent of the treatment components), the bar should be set high during training, as deterioration of skills is common post-training. With regard to the evaluation of the therapeutic process, several validated measures exist. For example, there are objective measures to assess whether counseling was consistent with a patient-centered communication (27,28). Strategies for assessing training to criterion are outlined in Table 3.

Monitoring skills over time

Training providers is conceptualized as an ongoing effort, rather than a one-time event. Ongoing coaching and feedback increases post-training proficiency and yields better treatment response (29). Booster training sessions and follow-up coaching should supplement regular supervision of providers. Offering continuing education credits, food, and other incentives helps to increase attendance at booster sessions. Regular supervision should occur with greater frequency immediately after training (e.g., weekly) and less often (bimonthly or monthly) once it is established that the training criterion has been maintained over time. Using a combination of supervision modalities is useful, such as group supervision, individual supervision, and peer-to-peer supervision. Audiotaping all encounters and randomly selecting some to listen to during supervision is optimal.

Immediately after training, it is recommended that a minimum of 50 percent of encounters are listened to (either during supervision or outside of supervision) in order to prevent protocol deviations. In the long-term, reviewing 20 percent of the encounters is optimal. If the individual falls below the a priori performance criterion,

then returning to 50 percent monitoring may be warranted. Feedback should be given in a supportive and constructive manner, to decrease defensiveness and increase learning. The learner's strengths should also be emphasized. Table 3 also shows several strategies to prevent provider drift over time.

Treatment delivery

Principles

The assessment and monitoring of treatment fidelity during treatment delivery involves treatment differentiation (did the providers only deliver the target treatment and not other treatments?), treatment competency (did providers maintain the skill set learned in training?), and treatment adherence (were the treatment components delivered as intended?) (11). This category is mutually exclusive from the above design and training categories, because well-trained providers may not always deliver a well-operationalized intervention protocol effectively, or with different participants across different contexts.

Treatments are less likely to be implemented with fidelity if they are complex, require many treatment providers, and use treatment manuals that are not user friendly. While it is unclear whether or not more experienced providers have higher treatment integrity, other provider factors such as acceptability of the treatment to the provider and providers' perceived effectiveness of the treatment have been shown to influence treatment implementation (25).

The usefulness of treatment manuals is controversial. On the one hand, they list the active treatment components and help to standardize treatments within and between providers. On the other hand, critics argue that they distance the patient from the provider, create passivity in the patient, and inhibit provider creativity. Kendall *et al.* (30) argue for a middle ground that does not compromise the fidelity of treatment, but, at the same time, calls for flexible adaptation which takes into account individual patient needs. For example, this could include administering treatment components out of order, dictated by the progression of the visit.

It is important to create relationships with providers so that they feel comfortable reporting treatment deviations. Integrity monitoring should be conducted in a collaborative versus hierarchical or critical manner. The rationale for monitoring should be explained to providers, as well as the implications of lack of treatment fidelity on the ultimate study outcome. It is also important to assess clinician beliefs and expectations about which treatment is more effective, and address these assumptions. The challenges of intervention delivery should be discussed with providers, and their ideas of how to improve integrity should be solicited (31).

Assessment of delivery

Both adherence to treatment components and competence to deliver the treatment in the manner specified (e.g. patient centered counseling) need to be assessed, as there are low correlations between the two behaviors (32,33). Nonspecific factors (e.g., empathy, communication style) should be assessed in order to minimize differences between providers and within providers over time. If there are significant differences between the groups and nonspecific factors are not assessed, it is difficult to conclude that the effects are due to the treatment rather than to different interactional styles. Differences between providers should be assessed through multiple methods on an ongoing basis, such as patient exit interviews, audiotaped sessions rated for nonspecific factors, and monitoring of participant complaints. Provision of feedback on interactional style should be given to providers.

The gold standard to ensure that treatments are delivered as specified is to use audiotapes or videotapes for objective verification of delivery, evaluated according to criteria developed a priori. Other methods of monitoring the fidelity of delivery, such as provider checklists (intervention checklists, encounter logs) and patient report (patient exit interviews), are less reliable and have low correlations with objective measures (34,35) but, nevertheless, can be used to supplement objective data. There are pros and cons of direct and indirect methods of monitoring (Table 1).

There are two purposes of assessment of fidelity of delivery: a) for use in supervision to improve provider skills and delivery, and b) for use in analytical models to determine the relationship between treatment fidelity and outcome. Monitoring for the latter purpose is more time-intensive, as it typically involves coding tapes. Monitoring for the former purpose is often more comprehensive, listening to the entire tape during supervision rather than just a portion (coding typically involves a truncated unit of analysis, such as a randomly chosen 20-minute segment). Regardless of the purpose of monitoring, all encounters are audiotaped and a random sample is chosen. Multiple sessions should be randomly selected from different phases of treatment. The provider should receive feedback on interactional style, treatment components omitted, treatment components added that were not specified by the protocol, dosage (number of minutes), and treatment differentiation (especially if the same provider is delivering different treatments).

If treatment fidelity data are being collected for inclusion in analytic models, or if an investigator is attempting to achieve provider consistency and standardization across multiple sites and supervisors, additional, more formal coding should occur. Raters of the audiotapes or videotapes should be independent of the study and blind to treatment assignment, participant progress and outcomes, and provider identity. In addition to achieving interrater reliability, raters

Table 4 Methods of Enhancing Treatment Fidelity: Treatment Delivery

-
- Create relationships with providers to increase their comfort for reporting deviations (collaborative versus hierarchical integrity monitoring).
 - Use a scripted curriculum or treatment manual.
 - Assess nonspecific effects through multiple methods and on an ongoing basis (patient exit interview, audiotape and code sessions, monitor participant complaints, provide feedback to provider).
 - Minimize differences within treatments and maximize differences between treatments: manuals, frequent supervision to catch mistakes early, limit contact between providers of different treatment conditions, monitor provider expectations about treatment.
 - Ensure adherence to the protocol (content, dose, and process): audio or videotaped encounters, provider self-monitoring and patient exit interviews.
 - Check for errors of commission and omission, degree to which treatment components were delivered, and nonspecific factors.
 - Establish minimum competency levels, below which providers are given remedial training (e.g., adherence to $\leq 80\%$ of the components).
 - Coders should be independent of the study, and blind to treatment assignment, participant progress and outcomes, and provider identity.
 - Use an independent group to review taped sessions and guess the treatment condition.
-

should also be skilled in treatment delivery as well as in more subtle aspects of the intervention and the treatment manual (32).

Several methods are used to code treatment fidelity data. The simplest method is to rate the occurrence or nonoccurrence of treatment components. A coder simply checks off the prescribed and proscribed components that occur while listening to the tape. A more detailed method is to rate the frequency of occurrence, degree of adherence to the component, and quality of delivery using Likert scales (e.g., 1 = none to 5 = very much).

Waltz *et al.* (9) recommends coding for the relative number of active treatment components versus inactive treatment components. Specifically, visits are coded for components that are a) unique and essential (i.e., components not found in the other approach being tested); b) essential to the treatment but not unique to it (i.e., empathy); c) compatible with the specified modality, and therefore are not prohibited but are neither necessary nor unique (chatting with client at the beginning of the session); and d) components that are proscribed. The proportion of observed versus possible components are computed. At least one other study has used this method with success (36). The disadvantage of the Waltz *et al.* (9) recommendation is that it is difficult to generate an a priori list of all of the proscribed elements, and there is a lack of clarity about what is essential but not unique.

A criterion for adherence to both nonspecific factors and to treatment components should be established. If providers do not achieve this criterion during treatment implementation, booster training sessions are recommended until the provider reaches the minimum level of competency that was established during the training. Competence or quality of delivery (e.g. communication skills) is distinct from provider adherence to treatment components, and both are predictive of treatment outcome (37). Shaw & Dobson (38) provided remedial training to providers who were rated on a validated measure as one standard deviation below their final training case. Although there is a lack of

clear guidelines about what the level of optimal level adherence should be, most agree that 80-100 percent integrity constitutes high fidelity, whereas 50 percent constitutes low fidelity (39,40). The strategies for assessing treatment fidelity during delivery are summarized in Table 4 and Appendix I.

Treatment receipt

Principles

Fidelity of treatment receipt refers to whether the treatment that was delivered to the participant was actually “received” by the participant. Treatment receipt involves whether or not the participant understood the treatment (as well as the accuracy of understanding), and demonstrates knowledge of, and ability to use, the skills or recommendations learned in the treatment. Checking on treatment receipt is especially important when participants are cognitively compromised or have low levels of literacy, education, or proficiency in English. If a patient does not understand or is not able to implement the new skills, then an otherwise perfectly designed and delivered intervention will not be effective. The strategies to enhance treatment receipt involve using methods to facilitate the participants’ comprehension of treatment (Table 5).

Assessment of treatment receipt

Assessment of treatment receipt involves verifying the participants’ understanding of the information provided in the treatment and verifying that they can use the skills and recommendations discussed (Appendix I). This could include written verification (pre–post tests), using audiovisuals (repeat information orally and visually), and behavioral strategies (role-plays skills with feedback). For example, in teaching parents how to brush their young child’s teeth, the parent could demonstrate the skills discussed during the visit. At the end of the encounter, the parent could be asked to rate their confidence that they could implement the behavior on a

Table 5 Methods of Enhancing Treatment Fidelity: Treatment Receipt

- Administer pre–post tests of client knowledge.
- Present material in engaging manner.
- Ensure that written materials have appropriate health literacy.
- Materials should be culturally relevant in terms of surface structure (photos) and deep structure (deeper cultural values).
- Provider should repeat information using multiple formats (verbal, pictures, written).
- Participant should be queried for their understanding of the material covered in the visit.
- Patients should role-play the skills and receive coaching and feedback.
- Assess patients' confidence to apply the skills delivered.
- Structure the intervention around achievement-based objectives.
- Collect and review self-monitoring data (e.g., brushing diary).
- Schedule follow-up visits and telephone calls to check in on understanding of the skills learned in treatment and level of adherence to recommendations.

1 to 10 scale (not at all confident to highly confident). If a parent says that they are an “x,” the provider could ask why they are at that number, and not a lower number. Then the provider could ask what could make the parent more confident so that they could implement the behavior (i.e., achieve a “10” on the scale).

Ensuring that audiovisual materials are culturally relevant is also paramount to treatment receipt. Cultural relevancy is enhanced by attending to surface structure (matching intervention materials and messages to the observable social and behavioral characteristics of the target population, such as people, places, language, music, foods, brand names, and clothing) and deep structure (incorporating the core cultural values of the target group to increase saliency of the message and program impact) (41). For example, changes in surface structure could include providing parents with handouts on substitutes for sugary snacks and sweets that list foods that are commonly consumed by that particular population. Changes in deep structure might incorporate faith or religion into some intervention materials or messages, for some groups. Other strategies to enhance treatment receipt are listed in Table 5.

Monitoring treatment receipt

A participant may be able to demonstrate understanding and ability to use the skill during the visit, but lose that understanding once they leave the office. Concepts and skills illustrated during visits can be further reinforced by the use of follow-up visits and phone calls. Goals could be set and challenges to goal implementation should be discussed. Participants can also self-monitor the target behavior (e.g., brushing child's teeth twice per day) using a calendar, recording the behavior, and noting the challenging times. Strategies that promote adherence as well as less effective strategies should be discussed.

Treatment enactment

Principles

Treatment enactment involves assessment, monitoring, and improving the ability of participants to perform treatment-related behavioral skills and cognitive strategies in relevant real-life settings (1,11). Treatment enactment is focused on whether skills are implemented in appropriate situations and at the appropriate time to have the intended effect on clinical and research outcomes (11). Enactment is an important addition to the treatment fidelity model because a distinction is made between what is actually taught (treatment delivery), what is learned (treatment receipt), and what is actually used (enactment) (11).

Enactment differs from the measurement of study outcomes because it is measured throughout the course of study implementation rather than only at the end of the study. Enactment is also different from patient adherence and treatment efficacy. In a dental health study, enactment entails visiting the dentist, adherence entails brushing the teeth using the recommended method, and efficacy is reduction of dental caries. In smoking cessation, enactment is buying the nicotine patch, adherence is using the patch, and efficacy is stopping smoking. Thus, it is possible to have a study with adequate enactment of treatment skills and poor treatment adherence or efficacy. If a study does not assess enactment, it is difficult to determine whether poor results are due to inadequate enactment or an ineffective intervention.

Assessment of enactment

Strategies for assessment include direct observation, self-report, and provider report. Enactment is usually assessed at a follow-up session or telephone call. This allows providers to assess and address the impediments to enactment. An example of an enactment checklist for oral health is listed in Table 6. Listed are the skills to be taught in the visit and practiced at home. These skills, although correlated with the outcome, are different from the actual outcome of cavity prevention.

Table 6 Example of Enactment Checklist for Oral Health

Visit 2		
Participant name:		
Observer name:		
Demonstrates proper brushing technique:	Yes	No
Demonstrates proper flossing technique:	Yes	No
Demonstrates knowledge about cavity prevention in children:	Yes	No
Demonstrates ability to reduce germs passed to baby:	Yes	No
Purchased beverages without sugar:	Yes	No
Purchased snacks with less sugar:	Yes	No
Refrained from putting baby to sleep with bottle in mouth:	Yes	No

A tool to assess treatment fidelity

My colleagues and I at the NIH BCC developed a questionnaire that allows investigators to assess the level of treatment fidelity in their own studies (1). The original version lists 25 treatment fidelity attributes that are rated as “Present,” “Absent, but should be present,” or “Not Applicable.” The measure contains items to assess the five categories of treatment fidelity (Design, Training, Delivery, Receipt, and Enactment). We used this measure to assess treatment fidelity across 10 years of health behavior change research (1). A total of 342 articles met inclusion criteria and were coded for their level of treatment fidelity. We found that 35 percent of studies used a treatment manual, 22 percent provided supervision for treatment providers, and 27 percent checked adherence to protocols. Only 12 percent used all three of these strategies and 54 percent used none of these strategies. The average proportion of adherence to treatment fidelity strategies in the Design category was 0.80, whereas the lowest mean proportion of adherence to strategies was in the Training category, where only 0.22 of strategies were reported. Delivery, Receipt, and Enactment categories were 0.33, 0.49, and 0.57, respectively. Only 15.5 percent of articles had 0.80 or greater proportion adherence to our checklist, across all categories. Appendix I displays an updated version of this checklist which contains more items focused on theory and on multicultural considerations. Investigators are encouraged to rate their own studies with the measure (both existing studies and those proposed in grants).

Our original measure was found to be reliable and valid (1,12). One study used our measure to assess treatment fidelity in 29 studies on secondhand smoke reduction. Studies with higher treatment fidelity ratings on our measure were more likely to obtain statistically significant results with an average fidelity rating of 0.74 for statistically significant studies versus 0.50 for statistically nonsignificant studies. After controlling for all relevant variables (year and location of study, efficacy versus effectiveness study, presence of theory, and intervention intensity), treatment fidelity as assessed by our measure was the only factor related to study outcome ($P = 0.052$). Three other studies provide working examples of the use our treatment fidelity measure in medical and community settings (15,42,43).

It has been recommended that the items on our measure should not be rated dichotomously, but should rather use a 5-point Likert scale (12). We had considered this during the development of our measure but believed that the subjectivity involved would make it difficult to proffer valid conclusions. In addition, a Likert scale would not enable an investigator to determine the “absent but should be present” category. A limitation of our measure is that it does not fully assess cultural relevancy (12), but this may be best addressed by a separate, more comprehensive measure that assesses all

of the nuances of cultural tailoring. There was also concern about the application of our treatment fidelity model (11,20) and measure (1) to real-world settings (44), although the measure was created from surveying the 15 BCC studies, all of which were hybrid efficacy-effectiveness studies. The ways in which our model and measure are applied to real-world settings are discussed in Resnick *et al.* (15).

Conclusion

Treatment fidelity enhances confidence in scientific findings, increases power to detect effects, and facilitates theory testing. Implementation of a treatment fidelity plan may require extra staff time and costs. However, the economic and scientific costs of lack of attention to treatment fidelity are far greater than the costs of treatment fidelity implementation. The model developed by the BCC and updated here outlines five, mutually exclusive domains of treatment fidelity. Lack of attention to any one domain heightens the risk of the inability to draw solid conclusions from the study.

This treatment fidelity model and accompanying measure is not meant to be a series of rigid steps but rather a set of guidelines to help investigators increase the likelihood of giving their treatments the fairest test possible. Flexible adaptation is called for within each of the domains. For example, study manuals need not be followed with such rigidity that the study’s hypotheses are actually undermined; training needs to be standardized but also flexibly adapted to different provider learning styles and levels of experience; treatment delivery needs to take into account different patient types and levels of motivation for change; treatment receipt must be tailored to the patient’s learning style and level of health literacy; and treatment enactment needs to be tailored to the person’s social and environmental context, as well as the economic and social contingencies that exist within that context. Flexible adaptation of interventions could also be gained by promotion of meta-competencies among providers, such as knowledge of research design, the goals of the study, and the underlying theory and rationale for the study. These strategies help to fulfill the goal of having fidelity with flexibility.

Acknowledgment

Grant support for this paper was provided by U54 DE019275-02, NIH/National Institute of Dental & Craniofacial Research.

Conflict of interest

The author declares no conflict of interest.

References

- Borrelli B, Sepinwall D, Ernst D, Bellg AJ, Czajkowski S, Breger R, DeFrancesco C, Levesque C, Sharp DL, Ogedegbe G, Resnick B, Orwig D. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J Consult Clin Psychol.* 2005;**73**(5): 852-60.
- Moncher FJ, Prinz RJ. Treatment fidelity in outcome studies. *Clin Psychol Rev.* 1991;**11**:247-66.
- Kazdin AE. Comparative outcome studies of psychotherapy: methodological issues and strategies. *J Consult Clin Psychol.* 1986;**54**(1):95-105.
- Yeaton WH, Sechrest L. Critical dimensions in the choice and maintenance of successful treatments: strength, integrity, and effectiveness. *J Consult Clin Psychol.* 1981;**49**(2):156-67.
- Lichstein KL, Riedel BW, Grieve R. Fair tests of clinical trials: a treatment implementation model. *Adv Behav Res Ther.* 1994;**16**:1-29.
- Henggeler SW, Melton GB, Brondino MJ, Scherer DG, Hanley JH. Multisystemic therapy with violent and chronic juvenile offenders and their families: the role of treatment fidelity in successful dissemination. *J Consult Clin Psychol.* 1997;**65**(5):821-33.
- DuFrene BA, Noell GH, Gilbertson DN, Duhon GJ. Monitoring implementation of reciprocal peer tutoring: identifying and intervening with students who do not maintain accurate implementation. *Sch Psych Rev.* 2005;**34**:74-86.
- Noel PE. The impact of therapeutic case management on participation in adolescent substance abuse treatment. *Am J Drug Alcohol Abuse.* 2006;**32**(3):311-27.
- Waltz J, Addis ME, Koerner K, Jacobson NS. Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *J Consult Clin Psychol.* 1993;**61**(4):620-30.
- Baer JS, Ball SA, Campbell BK, Miele GM, Schoener EP, Tracy K. Training and fidelity monitoring of behavioral interventions in multi-site addictions research. *Drug Alcohol Depend.* 2007;**87**(2-3):107-18.
- Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Ogedegbe G, Orwig D, Ernst D, Czajkowski S. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol.* 2004;**23**(5):443-51.
- Johnson-Kozlow M, Hovell MF, Rovniak LS, Sirikulvadhana L, Wahlgren DR, Zakarian JM. Fidelity issues in secondhand smoking interventions for children. *Nicotine Tob Res.* 2008;**10**(12):1677-90.
- Hansen WB, Graham JW, Wolkenstein BH, Rohrbach LA. Program integrity as a moderator of prevention program effectiveness: results for fifth-grade students in the adolescent alcohol prevention trial. *J Stud Alcohol.* 1991;**52**(6):568-79.
- Ellis DA, Naar-King S, Templin T, Frey MA, Cunningham PB. Improving health outcomes among youth with poorly controlled type I diabetes: the role of treatment fidelity in a randomized clinical trial of multisystemic therapy. *J Fam Psychol.* 2007;**21**(3):363-71.
- Resnick B, Bellg AJ, Borrelli B, DeFrancesco C, Breger R, Hecht J, Sharp DL, Levesque C, Orwig D, Ernst D, Ogedegbe G, Czajkowski S. Examples of implementation and evaluation of treatment fidelity in the BCC studies: where we are and where we need to go. *Ann Behav Med.* 2005;**29**(Suppl): 46-54.
- Zakarian JM, Hovell MF, Sandweiss RD, Hofstetter CR, Matt GE, Bernert JT, Pirkle J, Hammond SK. Behavioral counseling for reducing children's ETS exposure: implementation in community clinics. *Nicotine Tob Res.* 2004;**6**(6):1061-74.
- Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol.* 2008;**41**(3-4):327-50.
- McHugo GJ, Drake RE, Teague GB, Xie H. Fidelity to assertive community treatment and client outcomes in the New Hampshire dual disorders study. *Psychiatr Serv.* 1999;**50**(6): 818-24.
- Burns T, White I, Byford S, Fiander M, Creed F, Fahy T. Exposure to case management: relationships to patient characteristics and outcome. Report from the UK700 trial. *Br J Psychiatry.* 2002;**181**:236-41.
- Borrelli B, Resnick B, Bellg A, Ogedegbe G, Sepinwall D, Orwig D, Czajkowski S. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the Behavioral Change Consortium. *Seminar presented at the Annual Meeting of the Society of Behavioral Medicine*, Washington, DC, April 2002.
- Israel BA, Schulz AJ, Parker EA, Becker AB. Community-based participatory research: policy recommendations for promoting a partnership approach in health research. *Educ Health (Abingdon).* 2001;**14**(2): 182-97.
- Borrelli B. Smoking cessation: next steps for special populations research and innovative treatments. *J Consult Clin Psychol.* 2010;**78**(1):1-12.
- Kazdin AE. *Research design in clinical psychology.* 4th ed. Boston, MA: Allyn & Bacon; 2003.
- Crits-Christoph P, Mintz J. Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *J Consult Clin Psychol.* 1991;**59**(1):20-6.
- Borrelli B, Hecht JP, Papandonatos GD, Emmons KM, Tatevosian LR, Abrams DB. Smoking-cessation counseling in the home. Attitudes, beliefs, and behaviors of home healthcare nurses. *Am J Prev Med.* 2001;**21**(4):272-7.
- Roth AD, Pilling S. Using an evidence-based methodology to identify the competencies required to deliver effective cognitive and behavioral therapy for depression and anxiety disorders. *Behav Cogn Psychother.* 2008;**36**(2):129-47.
- Moyers TB, Martin T, Manuel JK, Hendrickson SM, Miller WR. Assessing competence in the use of motivational interviewing. *J Subst Abuse Treat.* 2005;**28**(1):19-26.

28. Ockene JK, Wheeler EV, Adams A, Hurley TG, Hebert J. Provider training for patient-centered alcohol counseling in a primary care setting. *Arch Intern Med.* 1997;**157**(20): 2334-41.
29. Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. *J Consult Clin Psychol.* 2004;**72**(6): 1050-62.
30. Kendall PC, Gosch E, Furr JM, Sood E. Flexibility within fidelity. *J Am Acad Child Adolesc Psychiatry.* 2008;**47**(9):987-93.
31. Power TJ, Blom-Hoffman J, Clarke AT, Riley-Tillman T, Kelleher C, Manz PH. Reconceptualizing intervention integrity: a partnership-based framework for linking research with practice. *Psychol Sch.* 2005;**42**(5):495-507.
32. Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: issues and recommendations. *Clin Psychol Sci Pract.* 2005;**12**:365-83.
33. Miller SJ, Binder JL. The effects of manual-based training on treatment fidelity and outcome: a review of the literature on adult individual psychotherapy. *Psychother Theory Res Pract Train.* 2002;**39**:184-98.
34. Carroll KM, Nich C, Sifry RL, Nuro KF, Frankforter TL, Ball SA, Fenton L, Rounsaville BJ. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug Alcohol Depend.* 2000;**57**(3): 225-38.
35. Wickstrom K, Jones K, LaFleur L, Witt J. An analysis of treatment integrity in school-based behavioral consultation. *Sch Psychol Quarterly.* 1998;**13**:141-54.
36. Collins SE, Eck S, Kick E, Schroter M, Torchalla I, Batra A. Implementation of a smoking cessation treatment integrity protocol: treatment discriminability, potency and manual adherence. *Addict Behav.* 2009;**34**(5):477-80.
37. Barber JP, Crits-Christoph P, Luborsky L. Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *J Consult Clin Psychol.* 1996;**64**(3): 619-22.
38. Shaw BF, Dobson KS. Competency judgments in the training and evaluation of psychotherapists. *J Consult Clin Psychol.* 1988;**56**(5):666-72.
39. Noell GH, Gresham FM, Gansle K. Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematical performance. *J Behav Educ.* 2002;**11**:51-67.
40. Holcombe A, Wolery M, Synder E. Effects of two levels of procedural fidelity with constant time delay on children's learning. *J Behav Educ.* 1994;**4**:49-73.
41. Resnicow K, Soler R, Braithwaite RL, Ahluwalia JS, Butler J. Cultural sensitivity in substance abuse prevention. *J Community Psychol.* 2000;**28**:271-90.
42. Spillane V, Byrne MC, Byrne M, Leatham CS, O'Malley M, Cupples ME. Monitoring treatment fidelity in a randomized controlled trial of a complex intervention. *J Adv Nurs.* 2007;**60**(3):343-52.
43. Wyatt G, Sikorskii A, Rahbar MH, Victorson D, Adams L. Intervention fidelity: aspects of complementary and alternative medicine research. *Cancer Nurs.* 2010;**33**(5): 331-42.
44. Leventhal H, Friedman MA. Does establishing fidelity of treatment help in understanding treatment efficacy? Comment on Bellg et al. (2004). *Health Psychol.* 2004;**23**(5): 452-6.

Appendix I Treatment fidelity assessment and implementation plan

Treatment Fidelity Strategies, Grouped by Category

Rate: Present, Absent but should be present, and Not Applicable. If present, describe the strategy used for that component.

Treatment Design

1. Provide information about treatment dose in the intervention condition
 - a) Length of contact (minutes)
 - b) Number of contacts
 - c) Content of treatment
 - d) Duration of contact over time
2. Provide information about treatment dose in the comparison condition
 - a) Length of contact (minutes)
 - b) Number of contacts
 - c) Content of treatment
 - d) Duration of contact over time
 - e) Method to ensure that dose is equivalent between conditions.¹
 - f) Method to ensure that dose is equivalent for participants within conditions¹
3. Specification of provider credentials that are needed.
4. Theoretical model upon which the intervention is based is clearly articulated.
 - a) The active ingredients are specified and incorporated into the intervention¹
 - b) Use of experts or protocol review group to determine whether the intervention protocol reflects the underlying theoretical model or clinical guidelines¹
 - c) Plan to ensure that the measures reflect the hypothesized theoretical constructs/mechanisms of action¹
5. Potential confounders that limit the ability to make conclusions at the end of the trial are identified.¹
6. Plan to address possible setbacks in implementation (i.e., backup systems or providers)¹
7. If more than one intervention is described, all described equally well.¹

Training Providers

1. Description of how providers will be trained (manual of training procedures)
2. Standardization of provider training (especially if multiple waves of training are needed for multiple groups of providers).
3. Assessment of provider skill acquisition.
4. Assessment and monitoring of provider skill maintenance over time
5. Characteristics being sought in a treatment provider are articulated a priori. Characteristics that should be avoided in a treatment provider are articulated a priori.
6. At the hiring stage, assessment of whether or not there is a good fit between the provider and the intervention (e.g., ensure that providers find the intervention acceptable, credible, and potentially efficacious¹
6. There is a training plan that takes into account trainees' different education and experience and learning styles.¹

Delivery of Treatment

1. Method to ensure that the content of the intervention is delivered as specified.
2. Method to ensure that the dose of the intervention is delivered as specified.
3. Mechanism to assess if the provider actually adhered to the intervention plan or in the case of computer delivered interventions, method to assess participants' contact with the information.
4. Assessment of nonspecific treatment effects.
5. Use of treatment manual.
6. There is a plan for the assessment of whether or not the active ingredients were delivered.¹
7. There is a plan for the assessment of whether or not proscribed components were delivered. (e.g., components that are unnecessary or unhelpful)¹
8. There is a plan for how will contamination between conditions be prevented¹
9. There is an a priori specification of treatment fidelity (e.g., providers adhere to delivering >80% of components).¹

Receipt of Treatment

1. There is an assessment of the degree to which participants understood the intervention.
2. There are specification of strategies that will be used to improve participant comprehension of the intervention.
3. The participants' ability to perform the intervention skills will be assessed during the intervention period.
4. A strategy will be used to improve subject performance of intervention skills during the intervention period.
5. Multicultural factors considered in the development and delivery of the intervention (e.g., provided in native language; protocol is consistent with the values of the target group).¹

Enactment of Treatment Skills

1. Participant performance of the intervention skills will be assessed in settings in which the intervention might be applied.
2. A strategy will be used to assess performance of the intervention skills in settings in which the intervention might be applied.

This checklist is adapted from: Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A.J., Czajkowski, S., Breger, R., DeFrancesco, C., Levesque, C., Sharp, D.S., Ogedegbe, G., Resnick, B., Orwig, D. (2005). A New Tool to Assess Treatment Fidelity and Evaluation of Treatment Fidelity Across Ten Years of Health Behavior Research. *Journal of Consulting and Clinical Psychology*, 73(5), 852-860.

¹Revisions made by B. Borrelli February, 2010.