



# Fidelity: An Essential Component of Evidence-Based Practice in Speech-Language Pathology

Joan N. Kaderavek

*University of Toledo, Toledo, OH*

Laura M. Justice

*The Ohio State University, Columbus*

**Purpose:** To provide a primer regarding treatment fidelity as it affects evidence-based practice (EBP) for speech-language pathologists.

**Method:** This tutorial defines treatment fidelity, examines the role of treatment fidelity for speech-language pathologists, provides examples of fidelity measurement, and describes approaches for assessing treatment fidelity.

**Conclusion:** Treatment fidelity is a neglected construct in the EBP literature; however, fidelity is a crucial construct for documenting intervention effectiveness and engaging in EBP.

**Key Words:** evidence-based practice, fidelity, intervention, clinical practice

The speech-language community is increasingly invested in identifying treatments with sufficient empirical support to justify their use within the context of evidence-based practice (EBP). EBP refers to a process in which one selects the most appropriate treatment (or assessment) for an individual. This process involves integrating one's knowledge of internal evidence (e.g., theoretical knowledge or child and family preferences) with one's knowledge of external evidence, including whether the treatment under consideration is empirically supported. The presumption is that outcomes for individuals with communication disorders are substantially improved when clinicians use treatments with a sound theoretical foundation, a systematic intervention delivery that can be achieved with fidelity, and a sufficient evidence base supporting efficacy or effectiveness (e.g., Heflin & Simpson, 1998; Justice & Fey, 2004; Simpson, 2005).

Within this tutorial, we discuss the construct of *treatment fidelity*, a sometimes-neglected construct in the EBP literature (Lane, Bocian, MacMillan, & Gresham, 2004). Treatment fidelity refers to the degree to which administration of a treatment corresponds to the prototype treatment, also referred to as the "gold standard" implementation (Hulleman & Cordray, 2009; O'Donnell, 2008). It is important for clinicians to understand what fidelity is and how it may be associated with treatment effects both within published research literature and within authentic (real-life) treatment applications. This tutorial is meant to provide a primer

regarding treatment fidelity and is organized to (a) define treatment fidelity as a construct relevant to research and practice, (b) examine the role and rationale of treatment fidelity and provide examples of fidelity measurements, and (c) describe approaches for assessing and improving treatment fidelity.

Throughout this tutorial, we use examples from our own research that has involved trials of language- and literacy-focused interventions within educational settings for children who exhibit risks in language and literacy development. Despite the focus on early childhood interventions within this article, we contend that the fidelity construct is one that is relevant to the field of speech-language pathology in all areas of clinical practice, including treatment of voice disorders, fluency, and adult language disorders.

## The Need for Documenting Fidelity: Clinical and Research Implications

Ensuring that the individual with a communication disorder receives scientifically based intervention is the "heart" of EBP. As part of the EBP paradigm, the speech-language pathologist (SLP) uses research evidence to help choose the best possible treatment, implements the intervention as described, and documents an individual's communication progress with carefully designed outcome measures (which may then lead to modifications of the intervention). In EBP, the outcome measures are the *dependent variable* (DV), as

these are dependent on the intervention being implemented. A DV also is the measure used to document change as a result of an intervention (i.e., “the experiment”). The SLP ensures reliability and validity of the DV by asking clinical questions such as *Will this outcome measure reliably capture changes in the individual’s performance over the course of intervention?* and *Does this outcome measure reflect the individual’s real-life communication ability?* Many SLPs are likely familiar with validity and reliability issues as they relate to outcome measures (Shipley & McAfee, 2009).

However, SLPs may less frequently consider fundamental issues related to the *independent variable* (IV). The IV is the intervention being implemented that, in turn, affects the DV. When SLPs think about intervention delivery, they confront the issue of fidelity. They ask clinical questions such as *Have I administered the intervention exactly as described?* and *What are the “active ingredients” of the intervention?* Both questions have clinical and research implications, because variation in IV fidelity affects clinical outcomes; both questions are discussed in the following sections.

### ***Have I Administered the Intervention Exactly as Described?***

**Background.** Fidelity is a component of the systematic process sometimes referred to as “scale up.” Scale up refers to the expansion of an intervention from a researcher-controlled intervention (i.e., implementation under ideal conditions or efficacy research) to evaluation of the treatment in natural settings whereby program developers and researchers have little or no face-to-face contact with the practitioners delivering the intervention (i.e., effectiveness studies). Ensuring the fidelity of the IV is central to the movement of an intervention through the scale-up process—beginning with small-scale pilot studies, to efficacy research, to effectiveness research.

The goal of *efficacy research* is to assess the causal relations between the IV and one or more clinically relevant outcomes (e.g., children’s vocabulary growth) in highly controlled settings. In contrast, *effectiveness research* (also called field research) is treatment-outcome research implemented within real-world settings. Effectiveness research evaluates whether treatment outcomes are generalizable across various settings and among various subgroups of recipients. Importantly, estimates of treatment effects are typically based on treatment implementation in which fidelity has been carefully controlled and rigorously monitored. In research, *treatment effect* refers to measure of the difference in outcome between intervention groups. In clinical work, treatment effect is shown when, for example, the SLP documents a student’s production of oral narrative prior to and following a period of intervention focusing on improving the student’s use of well-formed, narrative story structure.

The implications for fidelity measures vary at different levels of the scale-up process. During efficacy studies, the interventionist or researcher carefully identifies the essential ingredients fundamental to the intervention and sets a priori limits on acceptable levels of fidelity. Typically, treatment

implementation in efficacy studies is carefully designed to achieve 100% fidelity to the prototype. In fact, a critical quality indicator of an efficacy study is the assurance that the IV has been implemented with a high degree of fidelity (Gersten et al., 2004).

However, as an intervention is scaled up so as to examine its effectiveness, it is assumed that the fidelity of implementation will decrease as a result of contextual demands and individual variation. Similarly, when SLPs implement evidence-based research in the school or clinical setting, fidelity to the gold standard prototype is attenuated. Throughout the scale-up process, researchers and practitioners should evaluate the fidelity of implementation and consider the possible effects of fidelity variation.

As a consequence of fidelity variations, SLPs should consider two types of fidelity, intended fidelity and achieved fidelity. The two types differ based on the variation between treatment as it is intended to occur and treatment as it actually occurs. The *intended fidelity* of a treatment represents absolute fidelity to the gold standard treatment prototype; theoretically, this is the level of fidelity needed to achieve maximum treatment effects. In contrast, the *achieved fidelity* of a treatment represents the treatment as it is actually implemented.

**Clinical implications.** When implementing an intervention, the SLP tries to provide real-life treatment as similarly as possible to the prototype treatment. A close match between actual intervention and prototype treatment is described as *treatment integrity*. However, an actual treatment will always vary from the prototype treatment. *Treatment infidelity* represents the difference between intended fidelity and achieved fidelity. Treatment infidelity can be represented with the following notation (Hulleman & Cordray, 2009):

$$\begin{aligned} \text{Intended fidelity } (T^{\text{Tx}}) - \text{achieved fidelity } (t^{\text{Tx}}) \\ = \text{treatment infidelity } (T^{\text{Tx}} - t^{\text{Tx}}). \end{aligned}$$

With this formula, researchers document the variation between the close-to-100% fidelity as the intervention was conceived ( $T^{\text{Tx}}$ ) and the degree to which the interventionist varied from the gold standard prototype ( $T^{\text{Tx}} - t^{\text{Tx}}$ ). Although it is clear that researchers need to document the level of treatment infidelity in the scale-up process to verify a treatment’s level of effectiveness, the impact of treatment infidelity within real-life implementation is ambiguous. EBP research is not yet sophisticated enough to indicate the level of treatment infidelity that can be “tolerated” in clinical implementation. However, since we know that an individual’s positive communication outcome is dependent on an intervention’s “quality features” (i.e., features of the IV; Pence, Justice, & Wiggins, 2008; Ukrainetz, 2009), careful attention to fidelity is consistent with best practice (American Speech-Language-Hearing Association, 2005).

Low intervention fidelity, in fact, can negatively affect clinical decision making. Consider the clinical implications if an SLP implements a treatment with low levels of fidelity and the treatment does not result in predicted effects. In this case, the SLP does not know if the treatment was ineffective or if the treatment varied so much from the prototype

that results were attenuated. In the sections below, we will describe methods to maximize treatment fidelity.

In addition to comparing treatment fidelity to the prototype gold standard intervention (i.e.,  $T^{Tx}$ ), experts also consider treatment fidelity in relation to the counterfactual condition. The *counterfactual condition* describes what an individual would receive if the treatment had not been offered (Dane & Schneider, 1998; Hulleman & Cordray, 2009). Often in research studies, the counterfactual condition is described as “business as usual”; the business-as-usual condition represents what is being experienced by individuals who are not exposed to the experimental intervention.

SLPs consider the impact of the counterfactual condition when they work collaboratively with classroom teachers. A child’s exposure (or lack of exposure) to the IV within the counterfactual condition has a potential impact on the effectiveness of a communication intervention. Some classroom teachers already implement an intervention (e.g., provide targeted vocabulary enhancement or explicitly focus on phonological awareness skills). In comparison, another classroom teacher may rarely provide targeted interventions. The SLP modifies the frequency or type of SLP intervention in response to the child’s counterfactual exposure. One child may need only a brief exposure to the intervention to show positive effects. Another child may require intense exposure. Further, when SLPs read published EBP interventions, they should consider the impact of the counterfactual condition on data interpretation.

*Implications and examples from intervention research.* To illustrate the impact of treatment fidelity in a research study, consider a recent study in which researchers trained seven preschool teachers to implement a thematically organized language curriculum (Pence et al., 2008). As part of the curriculum, teachers were asked to implement dramatic play settings. Three times over a 9-month period, the researchers visited each teacher’s classroom and assessed fidelity to the curriculum components using a 45-item fidelity checklist. Twelve of the checklist items examined implementation of dramatic play settings; on average, teachers met 83% of these items across the three fidelity observations. Thus, the achieved fidelity ( $t^{Tx}$ ) was 83%. The level of treatment infidelity was calculated with the following formula:

$$100\% (T^{Tx}) - 83\% (t^{Tx}) = 17\% (T^{Tx} - t^{Tx}).$$

In other words, children in this study were exposed to 83% ( $t^{Tx}$ ) of the treatment components (with respect to the dramatic play features) and were not exposed to 17% ( $T^{Tx} - t^{Tx}$ ) of the components (treatment infidelity). It is unclear from this study’s findings whether this level of infidelity dampened the potential benefits of the curriculum to children’s language growth, as the relation between treatment fidelity and child gains attributable to intervention were not reported. However, research findings reported elsewhere have suggested that higher levels of treatment fidelity are associated with higher gains attributable to intervention (O’Donnell, 2008). In other words, treatment fidelity *moderates* (i.e., has an impact on) an intervention’s effectiveness (Justice, Mashburn, Pence, & Wiggins, 2008).

Research studies also demonstrate the effects of the counterfactual condition. To illustrate this concept, we refer back to the Pence et al. (2008) study. In this study, recall that the children in the experimental condition were exposed to 83% of the dramatic play components. However, since many children are exposed to dramatic play in their business-as-usual classroom, it is important to consider the counterfactual condition. Computing this formula documents children’s exposure to the targeted intervention above and beyond the counterfactual condition (i.e., the control condition). Using available data, the control teachers’ use of dramatic play components was assessed at three time points using the fidelity checklist in the same manner as occurred for the treatment teachers. Analyses revealed that the control teachers produced dramatic play components at a fidelity level of 36% across the three observations ( $t^C = 36\%$ ). Differentiation between the achieved fidelity of the experimental teachers and the fidelity of the control group teachers is computed as follows:

$$83\% (t^{Tx}) - 36\% (t^C) = 47\% (t^{Tx} - t^C).$$

This formula shows that there is a difference in treatment fidelity of 47% (i.e., difference between actual fidelity and fidelity of the control teachers = 47%). This can be interpreted to suggest that treatment teachers met about 50% more items on the fidelity checklist than teachers in the counterfactual condition. Importantly, this difference is statistically significant ( $p < .05$ ) and very large in size ( $d = 2.82$ ). This comparison highlights the point that treatment fidelity ( $t^{Tx}$ ) should be assessed not only in relation to intended fidelity ( $T^{Tx}$ ) but also to the counterfactual condition ( $t^C$ ). Understanding the variation between intended fidelity and the counterfactual condition is important for understanding the magnitude of effects an individual can be realistically expected to receive from a particular treatment.

### What Are the “Active Ingredients” of the Intervention?

*Background.* The goal of EBP is to (a) demonstrate that intervention (in general) has a positive impact, (b) demonstrate that some interventions are better than others, and (c) verify how and why a treatment is effective (Longabaugh et al., 2005). An intervention’s “how and why” are sometimes referred to as the *active ingredients*. As a field, speech-language pathology is increasingly focused on identifying the active ingredients behind specific intervention protocols.

*Clinical implications.* To operationalize an intervention, the practitioner implements an intervention’s active ingredients (Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005). These active ingredients may be based largely on theory, or they may be based on empirical evidence. Active ingredients most typically include specific treatment targets (e.g., specific grammatical forms), the therapeutic techniques (e.g., modeling these forms during interactive play), and the requirements for dosage (e.g., highly concentrated exposures several times per week). In combination, the active ingredients describe how and why the intervention brings about predicted outcomes.

The first set of active ingredients—identification of treatment targets and therapeutic techniques—is typically specified when an intervention is manualized. To increase fidelity, an intervention should have a treatment manual detailing specific behaviors to take place during the treatment (e.g., targets to be addressed, techniques and materials to be used, and expected behaviors of the participants). The treatment manual describes the gold standard of treatment implementation against which fidelity can be assessed. We provide more information about manualization of interventions in the following sections.

In addition to manualization, an intervention's active ingredients are articulated with identification of dosage. *Dosage* captures the frequency and duration of intervention administration (Warren, Fey, & Yoder, 2007). As part of EBP, practitioners consider a range of factors influencing dosage. Active ingredients with regard to dosage may include (a) the number of times the interventionist addresses a target or uses a technique during a given treatment session (e.g., 30 models in 30 min), (b) how long a treatment session should last, (c) how often treatment should be delivered throughout a week or month (e.g., 1 hr/week), and (d) the total length of required intervention across time (e.g., 9 months).

Fidelity measures will vary depending on the identified active ingredients. For example, if the recommended dosage is three times a week for 10 min/session, the practitioner can use simple logs to track the number of sessions an individual completes and the session length. If the dosage indicates that the number of targets/session must reach a threshold in order to be effective, then fidelity documentation becomes more fine-grained. In the section that follows, we make suggestions for several kinds of fidelity analyses.

*Implications and examples from intervention research.* As interventions move through the scale-up process, researchers typically will document an intervention's active ingredients as they occur in real life. Fidelity measures should document active ingredients relative to *procedure* (i.e., did the interventionist follow right steps and provide the appropriate dosage?) and *quality* (i.e., how well was the intervention delivered?).

Documentation of treatment procedure is generally straightforward; the researcher often uses a checklist to document fidelity. For example, Justice (2002) implemented a study to explore a vocabulary intervention. During the intervention, an adult provided 10 word-learning cues during storybook-reading interactions with preschool-age children. To assess procedural fidelity, 11 of 46 intervention sessions (24% of the corpus) were randomly selected and observed using the fidelity checklist. Fidelity to implementation protocols was high, ranging from 94% to 100% across the observed sessions, with a mean of 97%. Appendix A provides an example of a researcher-developed checklist used to assess procedural fidelity for the vocabulary intervention (Justice, 2002).

Documenting fidelity identifies moderators of an intervention's effectiveness. For example, the impact of treatment dosage is demonstrated in the authors' classroom-based language print-referencing intervention program. Specifically, Justice and colleagues (2008) examined levels of child attendance during a 30-week classroom book reading focusing on print referencing (see Justice, Kaderavek, Fan, Sofka, &

Hunt, 2009, for details of the intervention). Teachers' attendance logs for each child in the classroom provided an estimate of dose frequency. Data suggested that children's attendance levels moderated the effects of intervention on children's expressive language growth; fidelity documentation clarified the active ingredients.

Assessing fidelity relative to intervention quality is more difficult than documenting procedural fidelity. Assessment of treatment quality captures the manner in which a treatment is delivered (Mihalic, 2004). Qualitative measures document the dynamic processes of treatment essential to the intervention's active ingredients. For instance, measures of treatment quality may assess the sensitivity of the interventionist to the differing needs of participants receiving treatment, such as a child needing relatively high levels of scaffolding to acquire a particularly challenging concept. This component of fidelity assessment seeks to differentiate between treatments implemented well versus interventions implemented poorly. Since this is a more ambiguous process, qualitative fidelity often is not assessed (Carroll et al., 2007). Despite the challenges of operationalizing and assessing treatment quality, its assessment can be quite important, as a treatment can be implemented badly even when adherence to the procedure is high.

To highlight the need for assessing both procedure and quality as part of EBP, we refer again to the Pence et al. (2008) study. In this study, researchers documented treatment fidelity for 14 preschool teachers randomly assigned to implement (a) the Language-Focused Curriculum (LFC; Bunce, 1995) or (b) their current curriculum (e.g., Creative Curriculum; Dodge, Colker, Heroman, & Bickart, 2002; or High/Scope; Hohmann & Weikart, 2002). Fidelity was measured three times over the academic year using a 45-item fidelity checklist. Procedural fidelity was measured by documenting teachers' adherence to the step-by-step plans and activities specified in the LFC (e.g., reading a certain storybook or implementing a certain theme).

Qualitative aspects of intervention implementation were measured by documenting teachers' use of eight language stimulation techniques (LSTs) specified in the LFC. Stimulation techniques included *focused modeling*, whereby the adult contrasts two different language targets; *recasts*, in which the adult repeats the child's utterances using varied syntax; and *redirect/prompts*, whereby the adult prompts a child to interact with a peer. Trained observers spent approximately 2 hr in each classroom at the beginning of the school day and completed the 45-item treatment fidelity checklist while also collecting a video sample of the classroom interactions.

The analyses of the procedural and qualitative aspects of fidelity data showed two interesting results. First, as expected, the LFC teachers showed significantly higher levels of procedural fidelity to the LFC as compared with the non-LFC teachers. However, LFC teachers' procedural fidelity showed a pattern of high implementation of the treatment in the fall of the year, followed by a decrease in fidelity several months later.

Second, the LFC teachers and non-LFC teachers did *not* demonstrate significant differences in their use of qualitative features of the treatment. Despite the fact that the LFC



teachers had received a 3-day LFC professional training workshop and exhibited reasonably high levels of fidelity to the procedural features of the curriculum, the LFC teachers did not use LSTs significantly more than the non-LFC teachers. The authors concluded that “teachers are able to implement the more tangible aspects of a language curriculum with apparent ease, whereas the relational, dynamic aspects that require teachers to modify the way they converse with children pose a greater challenge” (Pence et al., 2008, p. 337). Importantly, research on the effectiveness of the LFC indicated that teachers’ use of LSTs was a significant moderator of children’s growth in expressive language skill. In sum, this study underscores the important role of fidelity data in interpreting data outcomes, as well as the need to assess both procedure and quality-related fidelity.

## Fidelity as an Emerging Construct in EBP

The importance of documenting treatment fidelity has been emphasized by experts both outside the field of speech-language pathology (e.g., Moncher & Prinz, 1991) and within the field (Ingham & Riley, 1998). Although literature on treatment fidelity appears across health science and education journals, in recent years it has been most closely considered within the fields of applied behavioral analysis, medicine, and mental health intervention (see Gresham, Gansle, & Noell, 1993; Gresham, Gansle, Noell, & Cohen, 1993; Peterson, Homer, & Wonderlich, 1982; Yeaton & Sechrest, 1981).

Despite the importance of treatment fidelity and its relation to treatment outcomes, most intervention research continues to fall short in this domain. For example, an examination of 152 school-based intervention studies reported in the *Journal of Applied Behavioral Analysis* between the years 1991 and 2005 demonstrated that nearly all (95%) provided an operational definition of the IV, yet only 30% provided treatment fidelity data. Not surprisingly, nearly half of the studies (45%) were judged to be at high risk for treatment inaccuracies (McIntyre, Gresham, DiGennaro, & Reed, 2007).

O’Donnell (2008) recently completed an analysis of fidelity data within the educational research literature. She analyzed research studies reporting results of core curriculum interventions and outcomes in kindergarten through 12th-grade classrooms. To be considered, the studies had to (a) be primary intervention research implemented by a single teacher (i.e., whole school programs were eliminated) and (b) examine the relationship between fidelity of implementation of core curriculum interventions and outcomes. Her analysis included published articles and unpublished doctoral dissertations from 1970 to 2008. Only 120 documents met the criteria for review, and of the initial pool of 120 only five studies met all established criteria for documentation of intervention fidelity. Fortunately, this trend is changing. Experts contend that fidelity measures are now a critical component of intervention research (Hulleman & Cordray, 2009; Troia, 1999).

Fidelity documentation is important even at the pilot study stages of EBP. At early study stages, fidelity data clarify important aspects of treatment execution, guide interpretation of external and internal validity, and facilitate study replication

and dissemination (Bruckenthal & Broderick, 2007). The value of fidelity measures at the pilot study level is demonstrated in a recent study. Girolametto, Weitzman, Lefebvre, and Greenberg (2007) evaluated the effectiveness of a 2-day training session for promoting the use of two emergent literacy strategies among early childhood educators. Sixteen early childhood educators were randomly assigned to an experimental or a control group. The in-service program for the experimental group focused on increasing educators’ use of abstract utterances and print references during shared book readings and a follow-up literacy activity (i.e., craft activity). The comparison group received training on facilitating peer interaction (The Hanen Centre, 2010). The reported results demonstrated that educators in the experimental program used more abstract utterances and print references during the follow-up literacy activity as compared to the comparison teachers. In addition, children in the experimental group responded more often with appropriate responses to abstract utterances and print references in comparison to children in the control group.

The researchers used fidelity analyses to elaborate and clarify their findings. The researchers reported that the individual fidelity profiles in the experimental group varied substantially among educators. In general, the print-referencing strategies were implemented more consistently than the use of abstract language. Fidelity data suggested the need for more training to attain consistency of delivery of the active ingredients. As this example demonstrates, fidelity helps understand salient issues even at early stages of intervention development.

Fidelity of implementation (or failure to ensure an intervention’s fidelity) can, in fact, have profound implications on the effectiveness of interventions as used at scale. An example of the negative outcome when fidelity is *not* established is evidenced in the implementation of the Drug Abuse Resistance (DARE) program (Rogers, 2003). Specifically, DARE was rapidly implemented in many schools in the 1980s and 1990s. One reason for its rapid adoption was the perceived flexibility of the intervention, such that components of the intervention could be readily altered or eliminated at the school level (O’Donnell, 2008; Rogers, 2003). This flexibility, however, runs counter to the principles of treatment fidelity. In the case of DARE, the program’s poor implementation fidelity as used in the field may have contributed to the U.S. Surgeon General of the United States’ statement in 2001 that DARE was unsuccessful (Satcher, 2001). Although it is impossible to tease apart the contribution of poor implementation fidelity, experts suggest lack of intervention fidelity likely affected DARE’s outcome (O’Donnell, 2008).

## Assessing and Improving Fidelity in Clinical Settings

### *Direct Versus Indirect Assessment of Fidelity*

Fidelity can be measured using both direct and indirect means (Cochrane, & Laux, 2008; Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000). Direct fidelity measures occur when the practitioner is directly observed either live or via a videorecording and fidelity is assessed using

some sort of objective observational tool with a priori coding categories. An example of a direct assessment is documented in Appendix B. This tool, the Fidelity Coding Catalog (FCC) represents a researcher-designed tool used to measure the use of print referencing during teacher-led book reading (Justice et al., 2009). The FCC documents the number of times a teacher addresses specific print-referencing targets identified for each book; the coding system is implemented when viewing teacher-led book group reading.

Direct observation results in the most thorough and objective data, whether one is assessing procedural or process intervention components. Negative aspects of direct observation include (a) the time and personnel requirement and (b) the fact that direct observation may not reflect the practitioner's "natural" implementation because he or she is aware of the observation (Cochrane & Laux, 2008). However, despite its challenges, direct observation is considered the gold standard.

Indirect fidelity measures are an alternative to direct assessment; indirect fidelity measures include self-report checklists and rating scales, interviews, logs, and permanent products (e.g., a client satisfaction survey and examples of student work following an educational intervention). Self-report checklists and rating scales allow practitioners to rate their compliance with targeted behaviors. In the following sections, we highlight both direct and indirect strategies to improve fidelity.

### ***Strategies to Increase Fidelity***

At this early stage in EBP implementation, the speech-language pathology literature does not document the effects of fidelity variations at the clinical level. So, in the following sections, our examples of strategies to improve SLPs' implementation fidelity are drawn mostly from our own laboratory research. We suggest that this tutorial highlights the need for clinical-level fidelity research.

**Manuals.** The treatment manual describes the salient features of the IV. Manuals can minimize the need for extensive training and often contain fidelity measures and outcome assessment strategies. Treatment manuals help make EBP accessible to providers and improve clinical decision making (Center for Substance Abuse Treatment, 2007).

The Section 1 Task Force of the Division of Clinical Psychology of the American Psychological Association discussed the importance of treatment manualization (Lonigan, Elbert, & Johnson, 1998). The task force stated that "manualized treatment not only allows statements to be made concerning specific interventions but also provides the necessary detailed description and standardization for replication, dissemination, and adequate training of [interventionists]" (Lonigan et al., 1998, p. 142). Treatment manuals can range from highly prescriptive lesson plans to a codified framework reflecting a general approach or set of principles for individualized implementation (Lonigan et al., 1998, p. 141).

As an example, in our research study (Justice et al., 2009), we trained experimental-group teachers to provide print-referencing behaviors during group book-reading sessions. Teachers read a different book each week for 30 weeks. The intervention's manualization had two components. First,

teachers received a manual describing the intervention's theoretical rationale and detailing the specific behaviors forming the intervention's active ingredients. Second, we placed a card describing two specific print-referencing targets in each book; the card gave examples of how the teacher should include the target within the book-reading session.

**Training.** Experts suggest that implementation of EBP presents both psychological challenges (e.g., resistance to change or commitment to current practices) and practical challenges, such as the additional training and supervision that may be required to implement evidence-based interventions with high fidelity (Center for Substance Abuse Treatment, 2007).

One of the goals of training is to make certain that interventionists follow the prototype procedure. An example from the health literature highlights the connection between manualization, training, and fidelity coding. Bruckenthal and Broderick (2007) documented the adherence of nurse practitioners delivering coping skills training for patients with chronic pain. The nurses' fidelity to the prototype treatment (as described in the treatment manual) was assessed during treatment sessions based on audiotapes listened to and coded by trained research staff. During the intervention, adherence to key training points was 86%. The lack of adherence was isolated to two main factors: (a) insufficient attention to ensuring that patients completed weekly homework assignments and (b) insufficient repetition/practice of patient pain-management skills. Importantly, these omissions helped explain variations in patient outcomes (i.e., self-efficacy and use of pain control coping measures). As a result of this outcome, the research team enhanced the training program to increase clinical fidelity.

Another example of training to increase treatment fidelity is demonstrated in our research (Justice et al., 2009). Participating teachers complete a 1-day training session prior to the school year and a 1-day refresher training session at mid-year. During the training sessions, in addition to the manual and a verbal overview of the 15 print-referencing targets, trainers provide videotaped illustrations of successful use of the print-referencing approach and provide opportunities for teachers to practice the intervention's active ingredients.

This training session was designed for classroom teachers who typically vary in their experience and educational levels. We hypothesize that SLPs—individuals with extensive training in delivering focused interventions—are likely to require less intense training to achieve high levels of fidelity. However, even with highly trained SLPs, in the future we envision more training opportunities (and perhaps training requirements) to increase intervention fidelity at the clinical level, particularly as more empirically validated treatments become available. It is likely that much of this training could be web-based. An optimal scenario would include web-based training with opportunities for practitioners to submit video clips for regular fidelity checks.

**Computer-assisted intervention.** Computer-assisted intervention is another means to increase intervention fidelity. Examples of computer-assisted intervention have been published, and their effectiveness has been documented (Shriberg, Kwiatkowski, & Synder, 1990). The role of the SLP can vary in computer-assisted intervention. In some cases, the software

“drives” the goal setting, stimuli exposure, and modifies intervention targets in response to the individual’s accuracy levels (e.g., Segers & Verhoeven, 2004).

In the future, it is likely that software may be used to guide intervention even within more traditional practitioner–client interactions. The software could become an online manual assisting the SLP to maintain high intervention fidelity. The software may potentially prompt the SLP to deliver specific interventions within specific dosage parameters and serve as a tracking device. The role of software to increase intervention fidelity is likely to be a focus of future intervention research.

*Documentation.* As we have previously discussed, fidelity should document both the procedure and the process of intervention; documentation of fidelity increases the consistency of implementation (Cochrane & Laux, 2008; Gresham et al., 2000; O’Donnell, 2008). Logs, check sheets, and patient surveys (i.e., questioning the individual receiving intervention about the components of the intervention) are viable fidelity tools. As an example of procedural fidelity, in our print-referencing research program, teachers completed an intervention log sheet each week for the 30-week intervention. Specifically, the teachers completed a checklist that identified whether they read the target book and whether they incorporated the designated print targets into the book reading. The teachers also documented any deviations in dosage (e.g., a day of school cancelled due to weather conditions). The teachers sent the intervention log sheet to the research team on a biweekly basis; the research team monitored the log sheets for fidelity, particularly as it related to dose. Research team members called or visited any teachers

with intervention deviations during the 30-week period to ensure accurate intervention implementation.

We provide other examples of direct and indirect fidelity measures in Table 1. The specific strategies used to document fidelity should be carefully selected and should be well aligned to the manualization of the treatment.

Direct observation, whether live or videotaped, is the preferred option for fidelity documentation. In our research program, every 2 weeks teachers would videotape themselves implementing a book-reading session and send the videotapes to research staff for analysis. The scheduling of the videotaping each week (i.e., taping on Monday, Tuesday, Wednesday, or Thursday) was randomized. The teachers sent their videotaped recording and 2 weeks of intervention log sheets in twice a month in preaddressed and stamped envelopes, resulting in 15 videotaped book readings submitted by each teacher.

The FCC fidelity analysis (see Appendix B) documented the number of times the teacher addressed the specific print-referencing targets identified for each book. The FCC involved a frequency count of each verbal reference to the target, coded as 1 = *yes*, 0 = *no*. Prior to completing the FCC teacher coding, each coder completed a self-study training module and reliably scored five training videos (approximately 10 min each) with 95% accuracy. After all tapes were coded, 10% of the videos were double-coded by a second trained coder. The training ensured that the fidelity coding would be completed with high reliability; interrater reliability was 91%.

FCC data were used to maintain each teacher’s adherence to the gold standard protocol. For this purpose, we provided written feedback to each teacher regarding intervention

**TABLE 1. Strategies for assessing and enhancing treatment fidelity.**

Focus	Strategies
Enhancing fidelity through practitioner training	<ul style="list-style-type: none"> <li>• Observe intervention implementation during role playing and according to an a priori checklist</li> <li>• Provide written exam pre- and posttraining</li> <li>• Videotape training in case there needs to be future training for other practitioners</li> <li>• Conduct regular “booster” training sessions</li> </ul>
Ensuring the active ingredients of the intervention	<ul style="list-style-type: none"> <li>• Provide a treatment manual</li> <li>• Monitor sessions via live observation or recordings and score practitioners on their adherence using a priori checklist in terms of process and procedure</li> <li>• Provide preset “packets” to trigger intervention (e.g., teachers get a specific protocol each week and are asked to repeat it a specified number of times per week)</li> <li>• Have practitioner(s) maintain log of implementation</li> <li>• Monitor subject output with regard to dosage (e.g., monitor homework)</li> <li>• Use scripted intervention protocol</li> <li>• Use computer prompts to trigger interventions</li> <li>• Have practitioners complete self-report questionnaires</li> <li>• Review tapes without knowing intervention condition and guess condition</li> <li>• After each encounter, have practitioner complete a behavioral checklist of intervention components delivered</li> <li>• Conduct participant interviews to determine whether certain intervention components were delivered</li> </ul>

Adapted from Bellg et al. (2004).

implementation as coded from two submitted videos. At Weeks 8 and 22, members of the research team reviewed the teacher videotapes and provided written feedback. Teachers who demonstrated less consistent fidelity were provided with follow-up visits and/or phone calls to clarify and reinforce the use of the print-referencing targets. Finally, fidelity data were documented and reported in published studies of the intervention's effectiveness.

*Sensitization to issues of fidelity when evaluating EBP.* One of the fundamental requirements to increase intervention fidelity is to increase awareness of the need for intervention fidelity. This process begins when practitioners read and evaluate intervention research. SLPs should consider fidelity measures when evaluating an intervention's published research evidence. Questions to ask when evaluating intervention research include the following:

- Does the researcher clearly describe the active ingredients of the IV?
- Does the researcher provide manualization and training to increase the fidelity of the IV?
- What procedures and tools (i.e., logs and coding of videotapes) were used to ensure IV fidelity?
- Does the researcher provide data documenting the fidelity of implementation of the intervention?
- Was the intervention implemented with high fidelity?
- If the fidelity of implementation varied, did the researcher account for fidelity variation as a moderating factor?

## Summary

In this article, we have highlighted the importance of treatment fidelity as a construct relevant to research and practice; we have also examined the role of treatment fidelity with regard to EBP. We have discussed issues in assessing treatment fidelity including manualization, training, coaching, and ensuring equivalency and dosage of treatment components. We illustrated how fidelity is documented in our own research and provide other examples to illustrate fidelity implementation.

As we have emphasized throughout this article, there are challenges associated with assessing fidelity and recognizing its role in understanding treatment effects. Interventions must be effective in real-world settings, but the procedures, training, manualization, dosage, and intervention quality must be consistently applied to produce optimal outcomes. Many practitioners will require continued training, guidance, and feedback to develop the skills to deliver the intervention as it was envisioned. To meet both demands, practitioners and researchers must collaborate and work together to achieve fidelity and flexibility (Dusenbury, Brannigan, Falco, & Hansen, 2003; Dusenbury & Hansen, 2004).

## Acknowledgments

Funding was provided by Grant R305G050005 from the U.S. Department of Education, Institute of Education Sciences (IES).

The content of this publication does not necessarily reflect the views or policies of the IES, nor does mention of trade names, commercial products, or organizations imply endorsement by the Department of Education. Portions of this article were presented at the 2007 Annual Convention of the American Speech-Language-Hearing Association, Boston, MA.

## References

- American Speech-Language-Hearing Association.** (2005). *Evidence-based practice in communication disorders* [Position statement]. Available from [www.asha.org/policy](http://www.asha.org/policy).
- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., ... Czajkowski, S.** (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology, 23*, 443–451.
- Bruckenthal, P., & Broderick, J.** (2007). Assessing treatment fidelity in pilot studies assist in designing clinical trials: An illustration from a nurse practitioner community-based intervention for pain. *Advances in Nursing Science, 30*, 72–84.
- Bunce, B.** (1995). *Building a language-focused curriculum for the preschool classroom. Vol. II: A planning guide*. Baltimore, MD: Brookes.
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S.** (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*, 40. Retrieved from [www.implementation-science.com/content/2/1/40](http://www.implementation-science.com/content/2/1/40).
- Center for Substance Abuse Treatment.** (2007). *Understanding evidence-based practices for co-occurring disorders. COCE Overview Paper 5* (DHHS Publication No. [SMA] 07-4278). Rockville, MD: Substance Abuse and Mental Health Services Administration and Center for Mental Health Services.
- Cochrane, W. S., & Laux, J. M.** (2008). A survey investigating school psychologists' measurement of treatment integrity in school-based interventions and their beliefs about its importance. *Psychology in the Schools, 45*, 499–507.
- Dane, A. V., & Schneider, B. H.** (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23–45.
- Dodge, D. T., Colker, L. J., Heroman, C., & Bickart, T. S.** (2002). *The creative curriculum for preschool*. Clifton Park, NY: Delmar Thomson Learning.
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W.** (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*, 237–256.
- Dusenbury, L., & Hansen, B.** (2004). Pursuing the course from research to practice. *Prevention Science, 5*, 55–59.
- Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, R.** (2005). *Implementation research: A synthesis of the literature* (FMHI Publication No. 231). Tampa: University of Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M.** (2004). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*, 149–164.
- Girolametto, L., Weitzman, E., Lefebvre, P., & Greenberg, J.** (2007). The effects of in-service education to promote emergent literacy in child care centers: A feasibility study. *Language, Speech, and Hearing Services in Schools, 38*, 72–83.
- Gresham, F. M., Gansle, K. A., & Noell, G. H.** (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis, 26*, 257–263.



- Gresham, F. M., Gansle, K. A., Noell, G. H., & Cohen, S. (1993). Treatment integrity of school-based behavioral intervention studies: 1908–1990. *School Psychology Review*, 22, 254–272.
- Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice*, 15, 198–205.
- The Hanen Centre. (2010). *Research summary: Learning language and loving it*. Retrieved from [www.hanen.org/web/Portals/0/HostedFiles/LLLRResearchSummary.pdf](http://www.hanen.org/web/Portals/0/HostedFiles/LLLRResearchSummary.pdf).
- Heflin, J. L., & Simpson, R. L. (1998). Interventions for children and youth with autism: Prudent choices in a world of exaggerated claims and empty promises. Part I: Intervention and treatment option review. *Focus on Autism and Other Developmental Disabilities*, 13, 194–211.
- Hohmann, M., & Weikart, D. P. (2002). *Educating young children: Active learning practices for preschool and child care programs*. Ypsilanti, MI: High/Scope Press.
- Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, 2, 88–110.
- Ingham, J. C., & Riley, G. (1998). Guidelines for documentation of treatment efficacy for young children who stutter. *Journal of Speech, Language, and Hearing Research*, 41, 753–770.
- Justice, L. M. (2002). Influences on preschoolers' novel word learning during shared storybook reading. *Reading Psychology*, 23, 87–106.
- Justice, L. M., & Fey, M. E. (2004, September 21). Evidence-based practice in schools: Integrating craft and theory with science and data. *The ASHA Leader*, 9(17), 4–5.
- Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers' early literacy development through teacher-child storybook reading. *Language, Speech, and Hearing Services in Schools*, 40, 67–85.
- Justice, L. M., Mashburn, A., Pence, K. L., & Wiggins, A. (2008). Experimental evaluation of a preschool language curriculum: Influence on children's expressive language skills. *Journal of Speech, Language, and Hearing Research*, 51, 983–1001.
- Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment integrity: An essential—but often forgotten—component of school-based interventions. *Preventing School Failure*, 48, 36–43.
- Longabaugh, R., Donovan, D. M., Karno, M. P., McCrady, B. S., Morgenstern, J., & Tonigan, J. S. (2005). Active ingredients: How and why evidence-based alcohol behavioral treatment interventions work. *Alcoholism: Clinical and Experimental Research*, 29, 235–247.
- Lonigan, C. J., Elbert, J. C., & Johnson, S. B. (1998). Empirically supported psychosocial interventions for children: An overview. *Journal of Clinical Child Psychology*, 27, 138–145.
- McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the *Journal of Applied Behavior Analysis* 1991–2005. *Journal of Applied Behavioral Analysis*, 40, 659–672.
- Mihalic, S. (2004). The importance of implementation fidelity. *Emotional & Behavioral Disorders in Youth*, 4, 83–86, 99–105.
- Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, 11, 247–266.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, 78, 33–84.
- Pence, K. L., Justice, L. M., & Wiggins, A. K. (2008). Pre-school teachers' fidelity in implementing a comprehensive language-rich curriculum. *Language, Speech, and Hearing Services in Schools*, 39, 329–341.
- Peterson, L., Horner, A., & Wonderlich, S. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis*, 15, 477–492.
- Rogers, E. (2003). *Diffusion of innovations*. New York, NY: Free Press.
- Satcher, D. (2001). *Youth violence: A report of the Surgeon General 2001*. Retrieved from [www.surgeongeneral.gov](http://www.surgeongeneral.gov).
- Segers, E., & Verhoeven, L. (2004). Computer-supported phonological awareness intervention for kindergarten children with specific language impairment. *Language, Speech, and Hearing Services in Schools*, 35, 229–239.
- Shipley, K. G., & McAfee, J. C. (2009). *Assessment in speech-language pathology: A resource manual* (4th ed.). Clifton Park, NY: Delmar.
- Shriberg, L. D., Kwiatkowski, J., & Snyder, T. (1990). Tabletop versus microcomputer-assisted speech management: Response evocation phase. *Journal of Speech and Hearing Disorders*, 55, 635–655.
- Simpson, R. L. (2005). Evidence-based practices and students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 20, 140–149.
- Troia, G. A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly*, 34, 28–52.
- Ukrainetz, T. (2009). Forward: How much is enough? The intensity evidence in language intervention. *Topics in Language Disorders*, 29, 291–293.
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, 13, 70–77.
- Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology*, 49, 156–167.

Received October 30, 2009

Revision received April 16, 2010

Accepted June 7, 2010

DOI: 10.1044/1058-0360(2010/09-0097)

Contact author: Joan Kaderavek, Department of Early Childhood, Special, and Physical Education (Mail Stop 954), University of Toledo, Toledo, OH 43606.  
E-mail: [joan.kaderavek@utoledo.edu](mailto:joan.kaderavek@utoledo.edu).

---

## Appendix A

### Procedural Fidelity Checklist for Vocabulary Intervention (Justice, 2002)

---

*Directions:* Place a score of 1 for items in which the condition was met. Place a score of 0 for items in which the condition was not met.

Score	Item
___	1. Child was seated facing the adult, with book open between adult and child (book facing child).
___	2. Book was read in its entirety.
___	3. Adult read all lines of text.
___	4. Adult responded to child's spontaneous comments with redirection back to the book.
___	5. Adult did not make extraneous comments (aside from cues on script) beyond the text until after book was read.
___	6. Implementation of vocabulary cue #1 per script.
___	7. Implementation of vocabulary cue #2 per script.
___	8. Implementation of vocabulary cue #3 per script.
___	9. Implementation of vocabulary cue #4 per script.
___	10. Implementation of vocabulary cue #5 per script.
___	11. Implementation of vocabulary cue #6 per script.
___	12. Implementation of vocabulary cue #7 per script.
___	13. Implementation of vocabulary cue #8 per script.
___	14. Implementation of vocabulary cue #9 per script.
___	15. Implementation of vocabulary cue #10 per script.

---

**FIDELITY SCORING:** Add all the numbers in the left-hand column to arrive at a raw score. Then, divide by the total number possible (15) to calculate a fidelity score.

Raw Score: \_\_\_\_\_/15

Fidelity Score: \_\_\_\_\_

---

## Appendix B

### Fidelity Score Sheet: Print-Referencing Book Reading

*Directions:* Score 1 each time a print-referencing comment occurs during the book reading.

	4 Print Domains (Key Words & Phrases)	Print-Referencing Comments (see manual for detailed list)	Tally	Raw Score
<b>WORDS</b>	Word(s)	<b>Words</b> (some words are short; others are long; letters make up words; words are distinct units of print, different than letters)  TEACHER COMMENTS:		
<b>LETTERS</b>	Letter (s) Uppercase, lowercase, capital Any letter name Any letter sound (with or without reference to letter name) ABCs, alphabet	<b>Letters</b> (letters come in 2 forms; names of different letters; the same letter can be used in many words)  TEACHER COMMENTS:		
<b>BOOK AND PRINT ORGANIZATION</b>	Read(s, ing) + discussion on book manipulation, discussion on directionality—for example, start(s, ing), stop(s, ping), end(s) Author, writer, written/wrote Title(d) Print Page(s) Book/story + called, named Book/story + poetry, nursery rhyme(ing), alphabet, fiction, nonfiction, information, real, pretend Book/story + front, cover, back, spine, beginning, end, part (i.e., any reference to structural features of books/parts) Book/story + upside-down, right-side up, turn, open, close (i.e., any reference to book manipulation)	<b>Book and Print Organization</b> (the order in which the pages are read; the job of the author; reading occurs from top of the page to the bottom; the role of the title; reading in English occurs from left to right; there are many different genres)  TEACHER COMMENTS:		
<b>PRINT MEANING</b>	Read(s/ing) NOTE: This serves as a stand-alone key word. "Read(s, ing)" should always be coded under Print Meaning, when not appropriately coded under Book and Print Organization (see above for coding definitions). Examples of a teacher's use of the key word "read(s,ing)": During discussion on a book's content (e.g., "Let's read this book to find out what it is about"), the "why" of reading (e.g., "We read to learn new things"), or use of a behavior management technique (e.g., "All eyes up here while I am reading").  Any reference to character or object in a book producing/containing words that are in print, as follows:  NOUN + VERB (such as <i>says, goes, has writing, cries, yells, tells, means, shows</i> + visible speech, visible sound, environmental print in illustrations (e.g., "Spot is talking here"; "The hat says/goes plop"; "The list has writing on it").  DO NOT code when pronouns are used, even when all other coding requirements are present.	<b>Print Meaning</b> (print function; words in the environment; discussion about the role and function of print and books for sharing information)  TEACHER COMMENTS:		
<b>Total Print-Referencing Comments:</b>				

#### Book-Reading Targets (Specific for Each Book of 30-Week Intervention)

During each book reading, the teacher is asked to include two specific print-referencing targets (per the card placed in the book). In this book reading:

Did the teacher include target #1? YES \_\_\_ NO \_\_\_

Did the teacher include target #2? YES \_\_\_ NO \_\_\_

TOTAL TARGETS INCLUDED\* (circle): 0 1 2

\*0 (no targets included), 1 (one target included), 2 (both targets included)