Treatment Integrity of School-Based Behavior Analytic Interventions: A Review of the Research

Kate E. Fiske, M.S., BCBA

Rutgers, The State University of New Jersey

ABSTRACT

A growing body of evidence suggests that treatment integrity of school-based behavior analytic interventions is related to intervention outcomes. These findings are of importance to behavior analysts, educators, and other practitioners working in school settings, and indicate that these professionals should be knowledgeable in the assessment of treatment integrity. In this article, we describe the methods used to measure treatment integrity in research and summarize the findings on consultation characteristics that affect treatment integrity. Based on the reviewed research, recommendations are offered to behavior analysts and school professionals to aid in the measurement and maintenance of treatment integrity in school settings.

Descriptors: School-based interventions, treatment integrity

Educational settings accommodate many children with learning disabilities, significant behavior problems, and developmental disabilities, guaranteeing a need for effective school-based interventions. Student response to intervention is the primary measure of the effectiveness of an intervention and the basis on which scientist-practitioners determine whether to modify, intensify, or terminate an intervention (Gresham, 2004). If student behavior improves following the implementation of an intervention, many assume the intervention is effective. In contrast, if student behavior worsens or remains stable, many assume the intervention is ineffective and seek to modify it. Response to intervention, however, is only a true measure of intervention effectiveness if the intervention is implemented as intended. The degree to which the treatment is implemented as planned is referred to as *treatment integrity* (Noell, Gresham, & Gansle, 2002).

For years, treatment integrity was considered only relevant for research-based treatment. It was—and still is—a primary methodological concern for researchers developing effective interventions. The pertinence of treatment integrity to the application of

school-based interventions for children with learning and behavior problems was largely overlooked and underestimated, but recent research has demonstrated the benefits of treatment integrity measures for school personnel. Despite these findings, the use of treatment integrity measures remains low even in practice-based research. For example, a recent analysis showed that 68% of published treatment studies did not measure the accuracy of implementation (Wheeler, Baggett, Fox, & Blevins, 2006). Such findings indicate that measures of treatment integrity may be rarely used in practice.

In this issue of *Behavior Analysis in Practice*, Vollmer, Sloman, and St. Peter Pipkin provided some practical guidelines for measuring and monitoring treatment integrity. This paper supplements that of Vollmer and colleagues by providing a review of research findings on treatment integrity, including factors that may influence treatment integrity and how best to maintain it over time, with a focus on school-based settings in particular. Although treatment integrity is relevant to interventions implemented at individual, group, and organizational levels, this paper focuses on its application to individual behavior analytic interventions.

We begin by examining the relation between treatment integrity data and treatment outcome in order to highlight the importance of treatment integrity in school-based interventions. We follow this with a discussion of issues related to measurement, training, and consultation. Finally, we offer recommendations for the assessment and strengthening of treatment integrity by behavior analysts and school personnel.

Treatment Integrity and Intervention Outcome

One of the most convincing arguments for maintaining high treatment integrity is that it strengthens the effects of intervention. Although the body of research in this area is limited, results support the commonly held assumption that higher treatment integrity results in better treatment effects (Arkoosh et al., 2007; Gresham, Gansle, Noell, Cohen, & Rosenblum, 1993; Holcombe, Wolery, & Snyder, 1994; Noell et al., 2002; Northup, Fisher, Kahng, Harrel, & Kurtz, 1997; Sterling-Turner, Watson, & Moore, 2002; Vollmer, Roane, Ringdahl, & Marcus, 1999; Wilder, Atwell, & Wine, 2006).

Experimental studies in which the integrity of an intervention is

systematically manipulated provide the most rigorous evaluation of this relationship. A handful of studies have examined the effects of treatment integrity on intervention outcome by implementing interventions with predetermined levels of integrity and evaluating the resulting effects on student behavior (Holcombe et al., 1994; Noell et al., 2002; Northup et al., 1997; Vollmer et al., 1999; Wilder et al., 2006). In several studies, researchers alternated varying levels of integrity of behavioral interventions and found greater improvements in behavior under conditions of high treatment integrity than under conditions of low treatment integrity (e.g., Holcombe et al., 1994; Noell et al., 2002; Wilder et al., 2006). Skills were also mastered more quickly when treatment was implemented with high integrity (Wilder et al., 2006).

It is important to note, however, that the effects of treatment integrity on outcomes often varied considerably across participants. For example, in a study examining the effects of treatment integrity on mathematics acquisition, students consistently performed well with perfect integrity (Noell et al., 2002). Surprisingly, however, some students performed just as well when the intervention was implemented with much lower treatment integrity. These data suggest low treatment integrity may not uniformly decrease outcomes for all students. Some elements of an intervention may remain robust despite low treatment integrity (Vollmer et al., 1999); in other cases, treatment components implemented with high treatment integrity may compensate for those implemented with low treatment integrity (Northup et al., 1997; Wilder et al., 2006). School professionals should not assume, however, that every intervention or child will sustain improved behavior with low levels of treatment integrity. Component analyses (i.e., systematic reductions of the integrity or presence of components) are necessary to identify components that maintain behavior change even when implemented with low integrity.

In sum, the current body of research suggests an emerging relationship between levels of treatment integrity and treatment outcomes. Knowledge of the measurement and consultation factors affecting treatment integrity will aid scientist-practitioners in choosing the best set of procedures to use in the classroom.

Measurement Factors Affecting Treatment Integrity

Several components of the measurement process, including the operational definition of the treatment and methods of measurement and calculation, may influence treatment integrity and the accuracy with which it is measured.

Operational Definition of Treatment

One crucial element when measuring treatment integrity is the operational definition of the intervention. All components of a behavioral treatment should be defined to promote high adherence to the treatment (Gresham, 1989) and encourage replication of the same treatment by multiple caregivers (Gresham et al., 1993). The behavioral interventions currently reported in many research studies are operationally defined on a molecular level, in which every step of the intervention is explicitly described. An example of such a definition includes the following: "Student will earn one sticker paired with verbal praise if he attempts to respond even if he responds with an incorrect answer" (DiGennaro, Martens, & Kleinmann, 2007, p. 450). By defining the intervention at a molecular level, behavior analysts may increase understanding of the intervention by those who implement it and permit a more objective assessment of treatment integrity. Thus, it is recommended that the treatment be defined for caregivers at this level.

Measurement of Treatment Integrity

The most frequently reported methods of treatment integrity measurement include direct observation, self-report, and permanent product

recording. Direct observation is a method in which a rater observes the implementation of the intervention *in vivo* and records the occurrence/nonoccurrence of elements of the intervention (see Vollmer et al., in this issue, for a detailed discussion of this method). Self-report measures require teachers to indicate through interviews or questionnaires the extent to which they implemented the intervention correctly. Permanent product recording is a measurement of the finished products of the intervention. For example, if a teacher was required to provide a child with feedback on his work, the presence of the teacher's correction marks on his worksheet would indicate that she successfully implemented this component of the intervention.

Direct observation can result in accurate measurement of many components, including those that cannot be measured through permanent products. However, it can also be time-consuming (Gresham, 1989). Another difficulty with direct observation is the possibility of teacher reactivity, in which the presence of the rater inflates the typical levels of integrity. For instance, Reid, Parsons, Green, and Schepis (1991) and Brackett, Reid, and Green (2007) showed that staff implementation of interventions increased when they were aware of the observation and evaluation of their behavior. In contrast, data collected during inconspicuous observations captured low levels of staff treatment integrity. Unscheduled, unobtrusive, and unannounced observation may best help reduce reactivity and increase accuracy of the measurement of treatment integrity.

Although self-report and permanent product measurement may be more efficient than direct observation, they also have limitations. Noell et al. (2005) demonstrated that teachers' self-reports of treatment integrity were unrelated to observational measures of treatment integrity and may inaccurately represent the true level of treatment integrity. Self-report is therefore considered the least reliable method of treatment integrity measurement (Noell, 2007).

Some authors also have speculated that permanent products may be created without implementing the intervention (Noell, Duhon, Gatti, & Connell, 2002), resulting in artificially high levels of treatment integrity. Perhaps even more challenging, however, not all interventions include components that produce permanent products (Noell, 2007).

As a result, a combination of methods may be best for collecting treatment integrity data in the classroom setting. Permanent products are extremely useful for collecting data on a daily basis and can be easily built into an intervention. However, observation of caregiver behavior should complement permanent product recording. Initially, frequent observations may be required following teacher training to monitor treatment integrity, but observations can be faded if treatment integrity remains high and stable. Both permanent product and daily, weekly, or bi-weekly observations lend themselves well to performance feedback, a method that is highly effective for increasing and maintaining teacher implementation of behavior intervention plans.

Calculation of Treatment Integrity

One of the challenges faced by behavior analysts collecting data on treatment integrity in classroom settings is identifying the best method with which to calculate treatment integrity data. In research, many methods are used. For interventions in which the teacher has the opportunity to implement each component only once during the observation, a simple method of dividing the number of components implemented correctly by the total number of components in the intervention, and multiplying by 100, has been sufficient (e.g., DiGennaro et al., 2007; DiGennaro, Martens, & McIntyre, 2005). The same method can be used with permanent product recording by dividing the number of permanent products produced by the teacher in a session by the total number of permanent products expected to

be produced during the session, and multiplying by 100 (e.g., Mortenson & Witt, 1998; Noell et al., 2000, 2005). In other studies, treatment integrity was calculated by dividing the number of intervals (e.g., 30 s) in an observation during which the teacher correctly implemented all components by the total number of intervals within the observation, and multiplying by 100 (e.g., Jones, Wickstrom, & Friman, 1997; Wood, Umbreit, Liaupsin, & Gresham, 2007).

When school professionals evaluate how to best collect treatment integrity data for interventions with multiple components that may be implemented multiple times within an observation, they should consider options that provide the most detail for individual treatment components. For example, Hagermoser Sanetti, Luiselli, and Handler (2007) examined the treatment integrity for a 27-component intervention designed to reduce problem behavior. During the observation, the observer collected treatment integrity data on each individual treatment component by rating whether the component was implemented as written for *every* opportunity during the observation or whether there were no opportunities to observe implementation of the component. Overall treatment integrity was then calculated by dividing the number of components scored as “implemented as written” by the total number of components observed. This approach provides detailed information on individual treatment components, allowing practitioners to identify components that are being implemented incorrectly so that specific feedback can be provided to teachers. In addition, this method allows practitioners to identify individual steps that may be crucial for improving child behavior.

Consultation Factors Affecting Treatment Integrity

Much discussion and surprisingly little empirical research has focused on the impact of consultation characteristics on the relation between treatment

integrity and intervention outcome. A number of authors have suggested characteristics (e.g., of the treatment, consultant, consultee, and environment) that may affect intervention implementation, but few have provided empirical support for these hypotheses (e.g., Dusenbury, Brannigan, Falco, & Hansen, 2003; Gresham, 1989; Noell, 2007). Consultation characteristics may be placed in two broad categories: those that are under the consultant’s control and those that are not under the consultant’s control.

Controllable Consultation Characteristics

Treatment characteristics (i.e., elements of the intervention) and consultant characteristics (i.e., the skills and behavior of the consultant) are largely under the consultant’s control. In addition, the training that consultants provide to caregivers is under the consultant’s control.

Treatment characteristics. The treatment characteristics most frequently hypothesized to impact treatment integrity are the complexity of the intervention and the time and materials required to implement it (e.g., Gresham, 1989; Yeaton & Sechrest, 1981). However, research findings on these factors have been inconsistent. Early research examining the relation between these characteristics and treatment integrity was based largely on teacher or caregiver reports of preference regarding these characteristics and not on the *actual* impact of complexity, time, or materials on treatment integrity (e.g., Witt, Martens, & Elliott, 1984). More recent research has disputed these early hypotheses. For example, Witt, Noell, LaFleur, & Mortenson (1997) found that the provision of materials required for intervention implementation did *not* significantly alter treatment integrity, as was originally hypothesized. Additionally, studies have shown rapid deterioration of treatment integrity following training, regardless of the complexity of the intervention (e.g., DiGennaro et al., 2007; Hagermoser Sanetti et al., 2007).

Consultant characteristics. Some researchers have debated whether consultants who work in collaboration with teachers produce more significant behavior change in the teacher than consultants who dictate intervention approaches without collaborating with teachers (e.g., Sheridan, 1992; Witt, 1990). Wickstrom, Jones, LaFleur, and Witt (1998) found the treatment integrity of interventions to be unaffected by the consultation approach used; integrity dropped quickly following training on the intervention across all teachers, regardless of whether a consultant collaborated with or prescribed an intervention to them. In fact, Noell (2007) suggested that a collaborative, co-investigator approach in consultation may even threaten treatment integrity by placing more emphasis on changing the behavior of the child rather than that of the teacher. Although the behavior of the child is a necessary focus, the behavior of the teacher in the context of treatment integrity should also receive attention.

Training and feedback. Research indicates that detailed and direct methods of staff training are most effective in improving initial implementation of an intervention (Dusenbury et al., 2003). For instance, Sterling-Turner, Watson, Wildmon, Watkins, and Little (2001) assessed the effectiveness of didactic training and direct modeling with performance feedback in teaching undergraduate students to implement a behavioral intervention. Modeling with feedback resulted in higher treatment integrity following training than did didactic instruction. Nonetheless, extensive training alone does not ensure prolonged high treatment integrity. In Mortenson and Witt (1998), for example, teachers demonstrated high levels of treatment integrity following an initial training comprised of component review, discussion of intervention rationale and importance, and *in vivo* training. Despite this intensive training, treatment integrity quickly dropped below desired levels within a week of intervention implementation.

To respond to this concern, much

research has focused on the effects of continued performance feedback on teachers' treatment implementation (Gilbertson, Witt, Singletary, & VanDerHeyden, 2007; Jones et al., 1997; Noell, Duhon, et al., 2002; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997; Noell et al., 2000, 2005; Witt et al., 1997). In a performance feedback model, consultants collect treatment integrity data through direct observation or permanent product recording and meet with the teacher to share data on treatment integrity and student behavior. Frequently, the data are graphed to allow for visual inspection of performance. Components incorrectly implemented during the session are identified and practiced, and the importance of missed components is reviewed. The consultant and the teacher problem-solve ways in which treatment integrity can be improved, and praise is provided for all components of the intervention that were implemented correctly (Noell et al., 1997).

To illustrate, Witt et al. (1997), Noell et al. (1997), and Jones et al. (1997) compared daily performance feedback to a more traditional model of classroom consultation. Traditional consultation (didactic instruction and provision of rationale for intervention) was ineffective in maintaining teacher implementation of individualized interventions. However, daily performance feedback was associated with 80% or higher treatment integrity levels in all studies. Consistently high integrity, however, was maintained only while performance feedback was in place; treatment integrity was more variable when feedback was discontinued in the maintenance phase.

Performance feedback is also effective when provided less frequently, such as once weekly (Mortenson & Witt, 1998) or even bi-weekly (Coddington, Feinberg, Dunn, & Pace, 2005). These less intensive consulting schedules allow for more prolonged provision of performance feedback, which results in greater maintenance of teacher skills (Coddington et al., 2005).

Given the impressive impact of performance feedback on treatment integrity, a number of studies have focused on the essential components of the performance feedback package. Findings indicate that consultant contact alone is not sufficient to increase treatment integrity and that data review is essential (Noell et al., 2000, 2005). The data review must focus on the teacher's performance rather than on student behavior (DiGennaro et al., 2007). Performance feedback sessions that focused only on student behavior or teacher commitment to the improvement of student behavior were not as effective as full performance feedback sessions (DiGennaro et al., 2007; Noell et al., 2005). Finally, graphing data on teacher behavior has been shown to be an essential component of performance feedback, as sessions conducted without visual graphs have been shown to be less effective in improving and maintaining treatment integrity (Noell, Duhon, et al., 2002; Hagermoser Sanetti et al., 2007).

The mechanism by which performance feedback impacts the behavior of teachers is as yet undetermined. The results are not merely an outcome of training, as integrity was typically high following training in most studies (e.g., DiGennaro et al., 2005; Mortenson & Witt, 1998; Noell et al., 2000). Performance feedback may maintain high integrity as a result of positive reinforcement that teachers receive for correct intervention implementation (i.e., praise and recognition by the consultant.) Alternatively, correct implementation may be negatively reinforced by the avoidance of feedback. Research has focused largely on the latter explanation, noting that performance feedback sessions often are only conducted if the teacher achieves less than perfect treatment integrity during an observation session (e.g., DiGennaro et al., 2005, 2007; Gilbertson et al., 2007). By maintaining high levels of treatment integrity, teachers avoid performance feedback sessions, which may include

such undesirable activities as discussion of errors, role play (DiGennaro et al., 2005), or directed rehearsal (Ward, Johnson, & Konukman, 1998) and can be time-consuming for the teacher. A study that examined this relationship failed to completely isolate the effects of negative reinforcement and positive reinforcement on teacher behavior (e.g., DiGennaro et al., 2005), and further research in this area is warranted.

Uncontrollable Consultation Characteristics

Several elements of the consultation process that may influence integrity appear beyond consultant control. These include consultee characteristics (e.g., beliefs, attitudes, experiences, and demographics of the treatment agent), client characteristics (e.g., characteristics of the student, including behavior), and environmental characteristics (e.g., the structure of the organization in which the intervention is implemented). For instance, research indicates that experienced teachers implement interventions more poorly than less experienced teachers (Noell et al., 2000). In addition, low integrity can be significantly predicted by poor student behavior (Hansen, Graham, Wolkenstein, & Rohrbach, 1991).

However, many assumptions in the literature about these uncontrollable characteristics are based largely on conjecture; the little empirical research that has emerged does not indicate that these characteristics affect treatment integrity. For example, teachers' attitudes about the acceptability of the proposed intervention, or the degree to which an individual perceives a treatment procedure as fair, reasonable, appropriate, and unintrusive (Kazdin, 1980), have been hypothesized to affect treatment integrity (Sterling-Turner & Watson, 2002). In contrast, Wickstrom et al. (1998) and Noell et al. (2005) found acceptability to be unrelated to reported integrity levels. Noell et al. also showed teacher acceptability to be unrelated to student outcome, indicating that some teachers found an

intervention acceptable even though the student's behavior did not improve.

In summary, results of this literature review suggest that most of the consultation characteristics hypothesized to affect treatment integrity are uncontrollable, too poorly studied to inform practice, or, if studied, have not demonstrated an effect on treatment integrity. In the face of such ambiguity, practitioners should focus on the one characteristic with substantial evidence for its effectiveness in increasing treatment integrity: the direct training and performance feedback provided to caregivers.

Conclusions and Guidelines for Best Practice

Emerging research findings have demonstrated a clear positive relationship between treatment integrity and intervention outcome. Methods used for measuring and evaluating integrity detailed in research provide excellent models for incorporating the assessment of treatment integrity into school-based interventions, and research evidence supporting the effectiveness of performance feedback provides school professionals with the tools needed to maintain treatment integrity. Nonetheless, research is still needed to better understand the impact of consultation characteristics on the maintenance of treatment integrity and to provide practitioners with direction in the face of potential barriers to treatment integrity. Current research indicates that performance feedback may overcome many of these barriers, but further investigation is needed to determine whether additional strategies may be even more effective and efficient. In the meantime, the following guidelines are recommended for monitoring and ensuring high levels of treatment integrity in school settings:

- *Recognize the positive relationship between treatment integrity and intervention outcome.*

High levels of treatment integrity have been demonstrated to result in more favorable intervention outcome than

low levels of treatment integrity. Although some students have shown improvement even with low treatment integrity, a component analysis of the intervention is necessary before school professionals can assume other students will exhibit the same resilience to low integrity in a classroom setting.

- *Provide clear, operational definitions of the intervention components.*

Definitions written at a molecular level, in which all components are explicitly described, may increase the understanding of an intervention and the objective assessment of treatment integrity.

- *Develop a reliable, accurate method for measuring treatment integrity.*

Unobtrusive, direct observation of treatment implementation is recommended. However, both permanent product and observational treatment integrity data can be combined and integrated into performance feedback sessions.

- *Choose a method for calculating integrity that will provide the most information about the individual treatment components.*

Methods that provide information about the implementation of each individual component will allow school professionals to provide specific feedback to teachers on their implementation of the intervention and to identify the components that are crucial to the effectiveness of the intervention.

- *Use detailed and direct training methods to teach staff.*

Modeling treatment components for staff and providing *in vivo* training whenever possible will maximize treatment integrity levels at the start of an intervention. Remember, however, that performance feedback is necessary to maintain high treatment integrity levels over time.

- *Provide on-going consultation consisting of performance feedback to maintain treatment integrity levels over time.*

Initially, daily feedback may be necessary to attain high levels of treatment integrity, but fading to weekly and bi-weekly feedback may result in greater

maintenance of treatment integrity. To maximize the effectiveness of performance feedback, feedback should always include data review that is focused on staff behavior rather than student behavior, praise for correct responding, and explicit correction for integrity errors; data review should include graphs of staff performance.

Although additional research is needed to further our understanding of the role of treatment integrity in school-based interventions, behavior analysts and other school professionals now stand to answer a call from the field to include treatment integrity measurement as an integral part of the behavioral intervention process. As such, scientist-practitioners should strive to implement the practices in treatment integrity measurement and maintenance described here to strengthen and maximize the effectiveness of behavior analytic school-based interventions.

References

- Arkoosh, M. K., Derby, K. M., Wacker, D. P., Berg, W., McLaughlin, T. F., & Barretto, A. (2007). A descriptive evaluation of long-term treatment integrity. *Behavior Modification, 31*, 880-895.
- Brackett, L., Reid, D. H., & Green, C. W. (2007). Effects of reactivity to observations on staff performance. *Journal of Applied Behavior Analysis, 40*, 191-195.
- Codding, R. S., Feinberg, A. B., Dunn, E. K., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis, 38*, 205-219.
- DiGennaro, F. D., Martens, B. K., & Kleinmann, A. E. (2007). A comparison of performance feedback procedures on teachers' treatment implementation integrity and students' inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis, 40*, 447-461.
- DiGennaro, F. D., Martens, B. K., & McIntyre, L. L. (2005). Increasing treatment integrity through negative reinforcement: Effects on teacher and student behavior. *School Psychology Review, 34*, 220-231.
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*, 237-256.
- Gilbertson, D., Witt, J. C., Singletary, L. L., & VanDerHeyden, A. (2007). Supporting teacher use of interventions: Effects of response dependent performance feedback on teacher implementation of a math intervention. *Journal of Behavioral Education, 16*, 311-326.
- Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review, 18*, 37-50.
- Gresham, F. M. (2004). Current status and future directions of school-based behavioral interventions. *School Psychology Review, 33*, 326-43.
- Gresham, F. M., Gansle, K. A., Noell, G. H., Cohen, S., & Rosenblum, S. (1993). Treatment integrity of school-based behavioral intervention studies: 1980-1990. *School Psychology Review, 22*, 254-272.
- Hagermoser Sanetti, L. M., Luiselli, J. K., & Handler, M. W. (2007). Effects of verbal and graphic performance feedback on behavior support plan implementation in a public elementary school. *Behavior Modification, 31*, 454-465.
- Hansen, W. B., Graham, J. W., Wolkenstein, B. H., & Rohrbach, L. A. (1991). Program integrity as a moderator of prevention program effectiveness: Results for fifth-grade students in the adolescent alcohol prevention trial. *Journal of Studies on Alcohol, 52*, 568-579.
- Holcombe, A., Wolery, M., & Snyder, E. (1994). Effects of two levels of procedural fidelity with constant time delay on children's learning. *Journal of Behavioral Education, 4*, 49-73.
- Jones, K. M., Wickstrom, K. F., & Friman, P. C. (1997). The effects of observational feedback on treatment integrity in school-based behavioral consultation. *School Psychology Quarterly, 12*, 316-326.
- Kazdin, A. E. (1980). Acceptability of alternative treatments for deviant child behavior. *Journal of Applied Behavior Analysis, 13*, 259-273.
- Mortenson, B. P. & Witt, J. C. (1998). The use of weekly performance feedback to increase teacher implementation of a prereferral academic intervention. *School Psychology Review, 27*, 613-627.
- Noell, G. H. (2007). Research examining the relationships among consultation process, treatment integrity, and outcomes. In W. P. Erchul & S. M. Sheridan (Eds.), *Handbook of research in school consultation: Empirical foundations for the field*. Mahwah, NJ: Erlbaum.
- Noell, G. H., Duhon, G. J., Gatti, S. L., & Connell, J. E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review, 31*, 217-234.
- Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. *Journal of Behavioral Education, 11*, 51-67.
- Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly, 12*, 77-88.
- Noell, G. H., Witt, J. C., LaFleur, L. H., Mortenson, B. P., Ranier, D. D., & LeVelle, J. (2000). Increasing intervention implementation in general education following consultation: A comparison of two follow-up strategies. *Journal of Applied Behavior Analysis, 33*, 271-284.

- Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., et al. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34*, 87-106.
- Northup, J., Fisher, W., Kahng, S. W., Harrel, R., & Kurtz, P. (1997). An assessment of the necessary strength of behavioral treatments for severe behavior problems. *Journal of Developmental and Physical Disabilities, 9*, 1-16.
- Reid, D. H., Parsons, M. B., Green, C. W., & Schepis, M. M. (1991). Evaluation of components of residential treatment by Medicaid ICF-MR surveys: A validity assessment. *Journal of Applied Behavior Analysis, 24*, 293-304.
- Sheridan, S. M. (1992). What do we mean when we say "collaboration"? *Journal of Educational and Psychological Consultation, 3*, 89-92.
- Sterling-Turner, H. E. & Watson, T. S. (2002). An analog investigation of the relationship between treatment acceptability and treatment integrity. *Journal of Behavioral Education, 11*, 39-50.
- Sterling-Turner, H. E., Watson, T. S., & Moore, J. W. (2002). The effects of direct training and treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly, 17*, 47-77.
- Sterling-Turner, H. E., Watson, T. S., Wildmon, M., Watkins, C., & Little, E. (2001). Investigating the relationship between training type and treatment integrity. *School Psychology Quarterly, 16*, 56-67.
- Vollmer, T. R., Roane, H. S., Ringdahl, J. E., & Marcus, B. A. (1999). Evaluating treatment challenges with differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis, 32*, 9-23.
- Ward, P., Johnson, M., & Konukman, F. (1998). Directed rehearsal and preservice teachers' performance of instructional behaviors. *Journal of Behavioral Education, 8*, 369-380.
- Wheeler, J. J., Baggett, B. A., Fox, J., & Blevins, L. (2006). Treatment integrity: A review of intervention studies conducted with children with autism. *Focus on Autism and Other Developmental Disabilities, 21*, 45-54.
- Wickstrom, K. F., Jones, K. M., LaFleur, L. H., & Witt, J. C. (1998). An analysis of treatment integrity in school-based behavioral consultation. *School Psychology Quarterly, 13*, 141-154.
- Wilder, D. A., Atwell, J., & Wine, B. (2006). The effects of varying levels of treatment integrity on child compliance during treatment with a three-step prompting procedure. *Journal of Applied Behavior Analysis, 39*, 369-373.
- Witt, J. C. (1990). Collaboration in school-based consultation: A myth in need of data. *Journal of Educational and Psychological Consultation, 2*, 367-370.
- Witt, J. C., Noell, G. H., LaFleur, L. H., & Mortenson, B. P. (1997). Teacher use of interventions in general education settings: Measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis, 30*, 693-696.
- Witt, J. C., Martens, B. K., & Elliott, S. N. (1984). Factors affecting teachers' judgments of the acceptability of behavioral interventions: Time involvement, behavior problem severity, and type of intervention. *Behavioral Disorders, 9*, 95-104.
- Wood, B. K., Umbreit, J., Liaupsin, C. J., & Gresham, F. M. (2007). A treatment integrity analysis of function-based intervention. *Education and Treatment of Children, 30*, 105-120.
- 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.

Author Note

Special thanks to Sandra Harris, Lara Delmolino, and Suzannah Ferraioli for their helpful comments on earlier versions of this manuscript.