

SUSTAINING HIGH FIDELITY OF TEACHER IMPLEMENTED EVIDENCE-
BASED PRACTICES: PERFORMANCE FEEDBACK WITH SELF-MONITORING

By

Regina M. Oliver

Dissertation

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in

Special Education

August, 2011

Nashville, Tennessee

Approved:

Joseph H. Wehby

Daniel J. Reschly

Kathleen L. Lane

Mark W. Lipsey

Mark R. Wolery

To my parents William W. Oliver, II and Gloria E. Oliver (Desreuisseau) who taught me the value of learning and always doing my best. Although you are both with God I have felt your love and support every step of the way.

I love you and miss you both.

ACKNOWLEDGEMENTS

This dissertation was supported in part by the National Comprehensive Center for Teacher Quality (S282B050051), Leadership Training in Emotional Disturbance (H325D070083) and the Melvyn I. Semmel Dissertation Research Award 2011.

I would like to acknowledge the support of my most prestigious dissertation committee and especially my advisor Joseph H. Wehby who showed me the value of research and always giving back to our schools. I have learned so much from each of my committee members and am honored to have their names listed as part of this work. To Mark Wolery I can't thank you enough for reaching out all those times on the fourth floor.

Without the love and support of my cohort, The Fab 14, I would not have been successful at Vanderbilt. Although "we're kind of a big deal," I am inspired and humbled by the amazing people I have grown to love and respect that shared the journey of this doctoral program with me. I want to especially acknowledge my lunch bunch members Jill McDaniel and Rachel Robertson whose group therapy sessions every Thursday kept me sane.

Finally, I want to thank my family and friends for their love and support, but especially my dear friend Karen Stigliano. Your love, guidance, and enduring devotion meant so much. You always believed in me and never let me give up. I love you my friend.

TABLE OF CONTENTS

	Page
ABSTRACT	vi
LIST OF TABLES	vii
LIST OF FIGURES	viii
Chapter	
I. INTRODUCTION	1
Treatment Integrity	2
Performance Feedback	4
Self-Monitoring	6
Theoretical Basis	7
Summary	9
Good Behavior Game	10
Purpose	12
II. METHOD	14
Participants	14
Setting	15
Materials	16
Measures	16
Treatment Fidelity	16
Observer Training	17
Interobserver Agreement	17
Procedural Fidelity Measures	18
Social Validity Measures	19
Experimental Design	19
Procedures	20
Baseline	20
Training on the Good Behavior Game	20
Good Behavior Game with Performance Feedback	21
Good Behavior Game with Self-monitoring	22
Maintenance	23
III. RESULTS	24
Formative Evaluation	24

Interobserver Agreement and Procedural Fidelity	24
Baseline.....	24
Good Behavior Game with Performance Feedback.....	27
Good Behavior Game with Self-monitoring.....	28
Maintenance.....	28
Summative Evaluation	29
Social Validity	30
IV. DISCUSSION.....	37
Limitations and Future Research	40
Conclusions.....	42
Appendix	
A. MATERIALS AND TREATMENT FIDELITY FORMS	43
B. PROCEDURAL FIDELITY FORMS.....	48
C. SOCIAL VALIDITY FORMS.....	50
REFERENCES	52

ABSTRACT

Improving student outcomes requires effective evidence-based instructional and behavioral practices, but successful adoption and sustainability of these practices by classroom teachers requires support. The purpose of this study was to examine the use of performance feedback as a strategy to generate high levels of treatment fidelity of the Good Behavior Game (Barrish, Saunders & Wolf, 1969) an evidence-based classwide behavior management practice. Secondly the purpose was to extend the literature by examining if this support can be maintained with the use of teacher self-monitoring of treatment fidelity. A multiple baseline design across four teachers was used to examine the purpose of the study. Results indicate a functional relation exists between initial training and performance feedback on treatment fidelity of the Good Behavior Game. Baseline levels of treatment fidelity shifted from zero to above 90% once the independent variable was applied with mean levels at 97% or above across teachers. In addition, results indicate a functional relation exists between use self-monitoring to maintain previous levels of treatment fidelity established with training and performance feedback. Social validity indicates favorable teacher ratings of the Good Behavior Game and self-monitoring. Limitations of the study with conclusions and future research are discussed.

LIST OF TABLES

Table	Page
1. Pre- and Post-test Social Validity Ratings of the Good Behavior Game and Self-monitoring	32
2. Pre-test Social Validity Ratings of the Good Behavior Game Based on the Number of Teachers' Ratings for each Question	33
3. Post-test Social Validity Ratings of the Good Behavior Game Based on the Number of Teachers' Ratings for each Question	34
4. Pre-test Social Validity Ratings of Self-Monitoring Based on the Number of Teachers' Ratings for each Question.....	35
5. Post-test Social Validity Ratings for Self-Monitoring Based on the Number of Teachers' Ratings for each Question.....	36

LIST OF FIGURES

Figure	Page
1. Performance Feedback and Teacher Self-Monitoring on Treatment Fidelity of the GBG	26

CHAPTER 1

INTRODUCTION

Improving student outcomes requires effective evidence-based instructional and behavioral practices. Researchers can design careful studies to identify and develop evidence-based practices (EBP), but these practices will only be as effective as their successful adoption by classroom teachers (The Evidence-Based Intervention Work Group, 2005). Moreover, the successful adoption of EBP is challenging given the multiple demands placed on school personnel. Unfortunately, accurate implementation typically does not occur (Noell, Witt, Gilbertson, Ranier, & Freeman, 1997; Witt, Noell, LaFleur, & Mortenson, 1997). Successful implementation requires teachers to implement EBP with the level of fidelity necessary to achieve the intended results established through research. Otherwise, intervention effects are likely to be attenuated (Dusenbury, Brannigan, Falco, & Hansen, 2003). Achieving and maintaining adequate levels of treatment fidelity therefore may be one way to facilitate implementation of evidence-based practices and achieve desired student outcomes (Hairrell et al., 2011; Smith Daunic, & Taylor, 2007).

The issue of teacher effectiveness and improving student outcomes is also directly related to the teacher-quality mandate of No Child Left Behind (NCLB; Goe, Bell & Little, 2008) which has produced considerable national debate and pressure on school systems to ensure students are provided instruction by highly qualified teachers (Berliner, 2005). This mandate specifies necessary certifications and credentialing required of teachers to be considered highly qualified. Credentialing alone is insufficient to ensure

teachers have the skills and competencies to deliver evidence-based practices and improve student outcomes (Angrist & Guryan, 2008; Rice, 2003; Wayne & Youngs, 2003). The current national movement is toward ensuring teachers are highly effective and therefore the focus has shifted from mere teacher credentialing to teacher effectiveness at producing student outcomes (Goe et al., 2008). Some states have instituted merit systems based on teacher effectiveness and performance (Florida Department of Education, 2007) or have used teacher evaluation systems to improve instructional practices (Goe et al., 2008, Mathers, Olivia, & Lane, 2008). Raising teacher quality and effectiveness necessitates procedures to support and shape teacher behaviors in the classroom. Improving the consistency and implementation of practices so full treatment integrity can be achieved is an important part of this process (Hairrell et al., 2011; Han & Weiss, 2005; Lane, Bocian, MacMillan, & Gresham, 2004). The purpose of this investigation is to evaluate the use of training and performance feedback to first generate high levels of treatment fidelity of a classwide management program while building teacher fluency of implementation. Secondly, the purpose is to investigate the use of teacher self-monitoring without performance feedback to maintain previous levels of treatment fidelity and as a way to maintain skills and practices.

Treatment Integrity

The accuracy and consistency with which an intervention is implemented as designed has been termed treatment fidelity or treatment integrity (Peterson, Homer, & Wonderlich, 1982). High levels of accuracy and consistency in the implementation of the independent variable are necessary to establish a functional relation (Gresham, 1989) because they strengthen the internal validity of the study. Strong accuracy and consistency also increase the believability of the outcomes (Gresham, 1989; Kennedy,

2005). The emphasis on treatment integrity is strongly warranted: In studies where researchers measure and report indicators of treatment integrity, larger effects are present when treatment integrity is high (Durlak & DuPre, 2008; Dusenbury et al., 2000; Gottfredson, Gottfredson, & Hybl, 1993; Telzrow, McNamara, & Hollinger, 2000). Unfortunately, treatment fidelity tends to be under measured and reported in the research literature despite attention to this issue for several years (Gresham, Gansle, & Noell, 1993; Hagermoser, Sanetti, Gritter, & Dobey, 2011; Lane et al., 2004; McIntyre, Gresham, DiGennaro, & Reed, 2007; Mooney, Epstein, Reid, & Nelson, 2003; Snell et al., 2010).

Because evidence-based practices tend to be more effective when treatment fidelity is high, researchers should establish procedures both to support high levels of initial treatment fidelity and identify methods for maintaining levels of treatment integrity over time (Sanetti & Kratochwill, 2008). Unfortunately, teachers' treatment fidelity tends to decline within ten days after teachers begin implementation (Hagermoser Sanetti & Kratochwill, 2009; Hagermoser Sanetti, Luiselli, & Handler, 2007; Noell et al., 1997; Mortenson & Witt, 1998; Witt et al., 1997). The tendency for treatment fidelity to decrease or drop after training and support is withdrawn is concerning given the importance of treatment fidelity to student outcomes.

Recently, researchers have been exploring methods by which treatment fidelity can be established and maintained. The most well documented process of improving treatment fidelity is performance feedback (Alvero, Bucklin, & Austin, 2001; Han & Weiss, 2008; Scheeler, Ruhl, & McAfee, 2004). Researchers have developed a systematic line of inquiry examining treatment fidelity using various forms of performance feedback to increase the fidelity of teacher-implemented EBP (Coddling, Livanis, Pace, & Vaca,

2008; Jones, Wickstrom, & Friman, 1997; Noell et al., 1997). In fact, research on using performance feedback to establish and maintain treatment fidelity has prompted recommendations for its use as a necessary component to sustain teacher implementation of school-based programs (Han & Weiss, 2005; Yeaton & Sechrest, 1981).

Performance Feedback

Performance feedback is a method to provide information or knowledge about skills and behaviors following performance to allow individuals to adjust their performance (Alvero et al., 2001; Mortenson & Witt, 1998; Sulzer-Azaroff & Mayer, 1991). It can take various forms but generally involves written or verbal feedback to an individual regarding performance (Alvero et al., 2001) and can range in complexity. Martens, Hiralall, and Bradley (1997) provided teachers with a simple note listing the number of praise statements, student behavior, and whether the goal was met or not, whereas Coddling, Feinberg, Dunn, and Pace (2005) used a more complex approach including one-on-one meetings with the teacher that included praise for correct steps and corrective feedback for missing or incomplete steps. Noell et al. (2005) further increased the comprehensiveness of performance feedback with graphs of teacher and student performance in feedback sessions. The nature and timing of performance feedback and its effectiveness has been the focus of much of the research on treatment fidelity.

Performance feedback has been used to increase treatment fidelity in multiple studies of teacher-implemented student interventions and classroom management plans (Coddling et al., 2005; Coddling et al., 2008; DiGennaro, Martens, & McIntyre, 2005; Gilbertson, Witt, Singletary, & VanDerHeyden, 2007; Mortenson & Witt, 1998; Noell, Duhon, Gatti, & Connell, 2002; Noell et al., 1997; Noell, Witt, Lynn, Lafleur, Mortenson, Ranier, & LeVelle, 2000; Noell et al., 2005; Reinke, Lewis-Palmer, &

Merrell, 2008; Witt et al., 1997). The frequency and immediacy of performance feedback in these studies included verbal prompting either immediately, within five or 10 min after observation, daily, weekly, or every other week. The complexity of performance feedback ranged from brief weekly interviews; weekly interviews with an emphasis on commitment to implementation; performance feedback with praise and corrective feedback; and performance feedback with praise, corrective feedback, graphic data on student or teacher performance, and a commitment from the teacher to continue implementation. Research from these studies indicates more immediate and comprehensive performance feedback produce greater outcomes in treatment fidelity (Alvero et al., 2001; Noell et al., 2002; Noell et al., 2005; Witt et al., 1997).

Performance feedback can produce high levels of treatment fidelity, but fading this support is also critical for long-term success. In applied settings, administrators are responsible for providing performance feedback, but few have the time to do this continuously. As a result, researchers have developed strategies for fading performance feedback while maintaining treatment fidelity. Researchers use pre-established performance criteria (e.g., 90% fidelity of implementation) to fade feedback from daily to every other day, and then every third day (Noell et al., 2000; Noell et al., 2002). Fading performance feedback from daily to weekly also has been examined and found to be successful (DiGennaro et al., 2005). Whether using pre-established criteria to fade support over a series of successive days, or fading the delivery schedule of performance feedback, it may take several weeks before support can be terminated. More research on feasible means to maintain treatment fidelity in applied settings such as the use of self-reports or self-monitoring is necessary to alleviate the time demands of performance feedback (Hagermoser Sanetti & Kratochwill, 2009).

Self-Monitoring

Although research to this point has not directly examined self-monitoring as a means to sustain treatment fidelity, it may be one less intense and more feasible method to maintain the effects of performance feedback on treatment. Self-monitoring has been widely used as an intervention in school settings to change a variety of student, and, to a lesser degree, teacher behaviors (Webber, Scheuermann, & McCall, 1993). Self-monitoring has been defined as supervising one's own professional practice or systematically recording one's own behavior (Gravina, Austin, Schoedtder, & Loewy, 2008; Kilbourne, 1991). Similarly, self-evaluation is the process of making judgments about one's own performance based on data (e.g., self-recorded) to change or improve practices (Keller, Brady, & Taylor, 2005). Shake (1986) suggested teacher self-monitoring improves teachers' awareness of their own practices and seems less threatening than conventional methods of teacher evaluation. Self-monitoring allows a teacher to evaluate a specific area and make a plan and subsequent improvements in practice. It may even be self-reinforcing to see improvements in behavior (Shake, 1986).

Self-monitoring has been used with many types of behaviors such as increasing teacher use of praise (Kalis, Vannest, & Parker, 2007; Keller et al., 2005; Sutherland & Wehby, 2001) and opportunities to respond (Sutherland & Wehby, 2001), increasing paraeducators prompts and responses for use of augmentative and alternative communication with students (Bingham, Spooner, & Browder, 2007), safe posture (Gravina et al., 2008), study behavior (Morgan, 1987), and increasing staff interaction with clients of a residential facility (Burg, Reid, & Lattimore, 1979). Self-monitoring used in these studies typically involved filling out a self-monitoring sheet, questionnaire, or checklist (e.g., Allinder & BeckBest, 1995; Allinder et al., 2000; Belfiore & Browder,

1992; Browder, Liberty, Heller, & D'Huyvetters, 1986). Other methods have included filling a board with colored dot stickers to monitor frequency of targeted behaviors (Burg et al., 1979), tallying simple frequency counts on a sheet of paper (Bingham et al., 2007; Kalis et al., 2007), or self-recording on audio tape and graphing target behaviors (Hoover & Carroll, 1987; Keller et al., 2005; Sutherland & Wehby, 2001).

Findings from the research on teacher self-monitoring generally indicate (1) using self-monitoring is superior to not using self-monitoring (Allinder et al., 2000); (2) using self-monitoring sheets, questionnaires, or checklists show improvements in teacher behavior (Allinder & BeckBest, 1995; Belfiore & Browder, 1992; Browder et al., 1986); (3) using self-recording on audio tape shows improvements in teacher behavior (Hoover & Carroll, 1987; Keller et al., 2005; Sutherland & Wehby, 2001); and (4) using simple tally marks of frequency counts on a sheet of paper does not indicate strong evidence for its use as a self-monitoring technique (Bingham et al., 2007; Kalis et al., 2007). Self-monitoring, therefore, appears to produce changes in teacher behavior and may be a viable method to fade performance feedback support while still maintaining high levels of treatment fidelity thereby improving student outcomes and teacher quality (Hairrell et al., 2001; Kissel, Whitman, & Reid, 1983; Petscher & Bailey, 2006).

Theoretical Basis

One approach to conceptualizing the theoretical basis for why performance feedback and self-monitoring may be effective strategies to obtain accurate implementation of an EBP and then maintain that accuracy over time is through the instructional hierarchy or learning stages theory (Haring & Eaton, 1978). Based on this theory, learning takes place in a series of hierarchical stages beginning with (a) acquisition; (b) fluency and maintenance; (c) generalization; and finally (d) adaptation

(Haring & Eaton, 1978). During each stage of learning, the learner not only progresses through a hierarchy of learning, but also requires different strategies to facilitate that learning process (Haring & Eaton, 1978). Each stage is predicated on the learner having achieved the previous stage. In other words, in order for fluency to occur, the learner must first acquire the skills. The first two stages of the learning hierarchy, acquisition and fluency/maintenance are most relevant to the current study and will be described next.

The first stage of the learning hierarchy is acquisition. Individuals are just beginning to perform the skills required and are not able to perform the learning task accurately or with a high level of reliability (Haring & Eaton, 1978). For acquisition of a skill to occur, modeling, practice, and corrective feedback are all effective strategies (Alberto & Troutman, 2009; Smith, 1981). These strategies have been used in a variety of research regarding teaching new skills from student academic skills (Cates, 2005; Mercer, Mercer & Pullen, 2011) to teacher instructional skills (Stichter, Lewis, Richter, Johnson, & Bradley, 2006). Initial teacher training of an EBP, therefore, should incorporate content information with modeling of the strategy by the trainer and subsequent practice by the teacher. As the teacher practices the trainer or consultant should also provide corrective feedback. Research suggests these approaches to teacher training will be beneficial to skill acquisition of an EBP (Stichter et al., 2006).

In the second stage of the learning hierarchy, fluency and maintenance, the learner is beginning to perform the learning task accurately (Idol, 1993; Kuhn & Stahl, 2003). As the learner continues to accurately perform the task, the task becomes automatic and does not require the amount of effort required during the acquisition and initial fluency stage (Kuhl & Stahl, 2003). When the learner has continued to perform the task accurately with automaticity over a period of time, maintenance occurs (Cooper, Heron, & Heward,

2007; Idol, 1993; Mercer et al., 2011). During this stage, the learner requires frequent opportunities to perform the task with corrective feedback and praise until the learner is able to perform the task and continues without assistance (Alberto & Troutman, 2009; Mercer et al., 2011). As the number of opportunities to perform the task accurately increases, the learner automatizes each step of the learning task and performs it without hesitation (Alberto & Troutman, 2009; Mercer et al., 2011). Performance feedback, provided daily, closely aligns with this stage of learning and is an appropriate strategy to create teacher fluency with an EBP as effective performance feedback incorporates corrective feedback and praise (Alvero et al., 2001; Noell et al., 1997). Continuing the fluent and automatic implementation of an EBP requires maintenance of the learning task. Although repeated use of performance feedback may be one method to maintain a learning task or EBP, it is also time consuming. Self-monitoring therefore may be an efficient method of continuing evaluation of task performance and provide reinforcement for teachers to continue (Shake, 1986).

Summary

Simple teacher awareness of effective practices does not necessarily translate into implementation. Teachers require training on evidence-based practices coupled with support during implementation to enhance the learning process (Han & Weiss, 2005). One question for researchers becomes how best to support teachers' implementation of EBP to achieve high levels of treatment fidelity (i.e., acquisition and fluency) thereby improving student outcomes and teacher quality. A second and perhaps more important question is how support can be faded while still maintaining high levels of treatment fidelity (i.e., maintenance).

Performance feedback is a researched-based approach to change teacher behavior and more specifically treatment fidelity but these changes do not always maintain over time and can be labor intensive for school personnel in applied settings. Self-monitoring, a less labor intensive approach, also has strong evidence for its use to change teacher behavior, but up to this point has not been used to maintain treatment fidelity. A combination of the two strategies may be beneficial. Self-monitoring after initial performance feedback and high level of treatment integrity may provide an approach that can feasibly be used to improve the implementation of evidence-based practices in the classroom (Hagermoser Sanetti & Kratochwill, 2009). However, current research has not sufficiently addressed the use of self-monitoring after performance feedback to maintain high levels of treatment fidelity specifically with classwide behavioral interventions.

The Good Behavior Game

Evaluating the utility of teacher performance feedback and self-monitoring to achieve and sustain treatment fidelity is only useful if the classwide behavioral intervention is also an evidence-based practice. Therefore, the Good Behavior Game (GBG; Barrish et al., 1969), a long-studied classwide behavioral intervention, has been selected as the EBP to evaluate the effectiveness of these elements. The GBG is a packaged intervention that promotes prosocial behavior by (a) explicitly teaching prosocial behavior and systematically reinforcing instances of behavior and (b) employing positive peer pressure through group contingencies thus focusing on positive reinforcement of behavior rather than punishment. The GBG is an example of a group contingency in which the class is split into two or more groups or teams who compete against each other. Criteria for winning the game, rules that need to be followed, and the reward for winning are established prior to beginning the game. If any team member is

observed to not be following the rules (e.g., out-of-seat, talking-out), a mark is placed on the board. Any team that meets the established criteria wins the game and the specified reinforcement. If a team does not win, they would continue working while the other team is rewarded for meeting their goal. In the seminal study on the GBG, Barrish, Saunders, and Wolf (1969) evaluated the impact of the GBG during a math and reading period using a combined multiple baseline with a reversal condition. Results indicated a dramatic decrease in disruptive classroom behaviors by 84.3% over all baseline and intervention phases (Barrish et al., 1969).

Since the initial study in 1969, the GBG has been systematically evaluated and found to be effective across a wide variety of populations of students, including children from preschool to 6th grade (Harris & Sherman, 1973; Medland & Stachnik, 1972; Warner, Miller, & Cohen, 1977; Bostow & Geiger, 1976; Darveaux, 1984; Johnson, Turner, & Konarski, 1978; Saigh & Umar, 1983; Swiezy, Matson, & Box, 1992) and settings, including the library (Fishbein & Wasik, 1981), with adolescents with emotional and behavioral disorders (Salend, Reynolds, & Coyle, 1989), with students at risk for placement in behavioral classrooms (Darveaux, 1984), and with students in rural Sudan (Saigh & Umar, 1983). Current research on the GBG continues to expand the evidence-base for its use in urban schools and urban cafeterias (Lannie, & McCurdy, 2007; McCurdy, Lannie, & Barnabas, 2009), kindergarten classrooms with modifications (McGoey, Schneider, Rezzetano, Prodan, & Tankersley, 2010), urban high school students (Kleinman & Saigh, 2011), and even primary classrooms in Spain (Ruiz-Olivares, Pino, & Herruzo, 2010).

Altogether, researchers have found the GBG produces positive proximal effects for target behaviors such as disruption, aggression, oppositional defiance, and attention

problems and potential distal effects with repeated exposure to the GBG. In fact, The GBG has been termed a “universal behavioral vaccine” (Embry, 2002), endorsed as a promising practice for discipline and violence prevention by the American Federation of Teachers (American Federation of Teachers, 2000), endorsed by the What Works Clearing House for reducing problem behavior in elementary schools (WWC; Epstein, Atkins, Cullinan, Kutash, & Weaver, 2008), and as an evidence-based practice by the Substance Abuse and Mental Health Services Administration (SAMHSA; www.samhsa.gov). Moreover, the applied nature of the research on the GBG makes it an ecologically valid approach as it has been researched in actual classrooms with indigenous teachers and used to address behaviors of importance to teachers and other school personnel. As an ecologically valid evidence-based practice, the GBG is an example of a multi-component classwide intervention teachers may be asked to implement in their classrooms making it a relevant classwide practice to examine how training with performance feedback impacts treatment fidelity and whether this support can be successfully faded using teacher self-monitoring.

Purpose

The purpose of the this study was to examine the combined use of teacher performance feedback and teacher self-monitoring as a strategy to first generate high levels of treatment fidelity of an EBP and second fade this support while maintaining high treatment fidelity. The “Good Behavior Game” (Barrish et al., 1969) was selected as the classroom management practice based on numerous studies of its efficacy as a classroom management strategy (Darveaux, 1984; Dolan et al., 1993; Fishbein & Wasik, 1981; Harris & Sherman, 1973; Ialongo, Werthamer, Kellam, Brown, Wang, & Lin, 1999; Kellam, Ling, Merisca, Hendricks brown, & Ialongo, 1998; Lannie & McCurdy,

2007; Medland & Stachnik, 1972; Tankersley, 1995; van Lier, Muthén, van der Sar, & Crijnen, 2004). A multiple baseline across teachers design was used to address the following research questions:

1. What effect does didactic training of the Good Behavior Game have on initial levels of treatment fidelity?
2. What effect does performance feedback have on treatment fidelity of the GBG?
3. What effect does teacher self-monitoring without performance feedback have on previous levels of treatment fidelity with performance feedback?
4. How will teachers rate the Good Behavior Game and self-monitoring on social validity measure before and after the study?

CHAPTER II

METHOD

Participants

Participants were recruited through an initial phone call to school principals describing the study and requesting the principal to notify 2nd or 3rd grade teachers of the opportunity to participate in the study. The authors met with interested teachers to describe the study and obtain consent. Participants were not targeted or excluded based on race, gender, or ethnicity. Participants were four female, 2nd grade general education teachers from the same elementary school. Teachers ranged in age from 25 to 36 years ($M = 30.75$; $SD = 4.57$) and had 3 to 10 years teaching experience ($M = 6.5$; $SD = 3.5$) in the public schools. The highest level of education obtained for each teacher respectively was a bachelor's degree ($n = 1$), hours beyond a bachelor's degree ($n = 1$), and a master's degree ($n = 2$). All teachers were certified to teach elementary with one teacher also certified to teach English as a second language.

All participants were screened prior to the study for inclusion using a questionnaire with five questions the participants responded to on a 5 point Likert scale (1 = *Strongly Disagree*; 5 = *Strongly Agree*). To be included in the study, participants were required to teach either 2nd or 3rd grade and to answer *Agree* or *Strongly Agree* to questions to the following questions: (1) I have high levels of disruptive student behavior in my classroom; (2) The disruptive classroom behavior interferes with student learning; (3) I would like to decrease disruptive behavior and increase student time on task; (4) I would like assistance to improve student behavior in my classroom; and (5) I am willing

to try a new strategy. Two teachers responded *Strongly Agree* to all questions and all four teachers responded *Strongly Agree* to the question “I would like to decrease disruptive classroom behavior and increase student time on task” ($M = 4.7$).

Setting

The study took place in an elementary school in a large metropolitan school district located in the southern part of the United States. The school housed pre-k through 4th grade with a total of 550 enrolled students. The demographics of the school included 46.3% female, 2.7% Asian, 40.5% Black, 12.5% Hispanic, 0.6% American Indian, and 43.8% White. More than half of the children (67.7%) were eligible for free/reduced lunch and 8.4% had English as a second language. The summer prior to the 2010-2011 school year all teachers and Principals met to establish school-wide behavioral expectations although the only evidence of this was a poster of the expectations observed in one teacher’s classroom. Although not measured, previous experience conducting research in this school suggests a culture of strong leadership, particularly from the Assistant Principal. All sessions occurred in general education classrooms during language arts, a time the teachers designated as having the highest levels of disruptive behavior. The classrooms for Teachers 1, 2, and 4 were located within the main school building. The classroom for Teacher 3 was a portable trailer located just outside the main building. The number of students in the classrooms ranged from 19 to 20. *Treasures Series* by McMillian/McGraw-Hill plus supplemental materials was used by all four participants for the language arts curriculum. This curriculum is designed to provide a combination of explicit instruction, leveled literature, and multiple opportunities for students to practice skills.

Materials

Participants were provided a 4 page handout of the training slides describing the rationale, research, and components of the GBG. Participants also were provided an 80 by 50 cm laminated board to post in front of the classroom for use during the GBG sessions. Participants used the scoring board to write the team names, behavioral criteria, and tally mark next to the team name when a team member exhibited a rule violation. Participants were provided a 20.32 by 10.16 cm plastic pencil box to keep strips of paper with rewards written. Participants wrote potential rewards on strips of paper and stored them in the box. After the conclusion of the game, the participant would pull a strip of paper to announce the reward. A digital timer was provided to each participant to begin and end the game. A treatment fidelity checklist was used to collect data on the dependent variable during all phases including baseline. During the self-monitoring phase, participants were provided with several copies of a self-monitoring sheet. Examples of these materials and data collection sheets can be found in Appendix A.

Measures

Treatment fidelity. The dependent variable collected and used to make phase change decisions was percent treatment fidelity of steps implemented for the GBG. Data were collected through classroom observations by the consultant and a trained researcher using a treatment fidelity checklist and graphed daily. The first author was the consultant conducting all performance feedback sessions and was a doctoral candidate with more than 10 years' experience in schools. Researchers conducting observations were two advanced doctoral students with several years' experience working in classrooms. Treatment fidelity was defined in two ways. First, treatment fidelity quantity was defined as the number of completed steps divided by the total number of steps possible,

multiplied by 100 to obtain the percent of treatment fidelity observed. Second, treatment fidelity quality was defined through a rating system from 1 to 5 with lower numbers indicating no fidelity or low fidelity (1 = *not implemented*; 5 = *full fidelity*).

Observer training. Researchers were trained using three separate videos of GBG sessions in classroom settings. Researchers watched each video and scored treatment fidelity quantity and quality on the GBG treatment fidelity form separately for each video (Appendix A). The treatment fidelity scores were compared to the master scoring sheet to determine inter-observer agreement for both treatment fidelity quantity and treatment fidelity quality. Agreement for treatment quantity was calculated using the formula $(\text{agreements} / \text{agreements} + \text{disagreements} \times 100)$ to obtain point by point agreement percentage (Ayres & Gast, 2010). The mean agreement across observers was 99%. Treatment fidelity quality was calculated using the same formula as treatment fidelity quantity with two definitions of agreement. First, if quality ratings between consultant and researcher were within one point (e.g., 4 or 5), it was considered an agreement. The mean treatment quality agreement within one point across consultant and researcher was 97% (range 93 – 100%). Second, consultant and researcher were required to obtain the same rating in order for an agreement to occur. The mean exact agreement across raters for treatment fidelity quality was 92% (range 86 – 95%).

Interobserver Agreement

Interobserver agreement between the consultant and one of the researchers was collected during baseline and treatment phases across participants for 40% of the observational sessions. Interobserver agreement for treatment fidelity was collected by the consultant and the researcher simultaneously and independently using the GBG treatment fidelity checklist. Two IOA measures were calculated for treatment fidelity (1)

number of steps implemented and (2) quality rating for each step. IOA for the number of steps implemented was calculated as the number of agreements plus the number of disagreements divided by the total and then multiplied by 100 to obtain a percentage. Scores for the quality ratings were considered as an agreement if the consultant and researcher were within one rating of each other (e.g., consultant rated 5 but researcher rated 4). Agreement was calculated exactly as in the case for the number of steps implemented to obtain a percentage.

Procedural Fidelity Measures

Procedural fidelity was collected for performance feedback sessions between the consultant and researcher during GBG phases (see Appendix B). The researcher observed the performance feedback sessions between the consultant and teacher and checked that each step was performed. Specifically, the researcher checked that the consultant (a) listed each step that was performed by the teacher with the appropriate quality rating and provided praise for at least one; (b) described any steps that were missing or had a quality rating less than four and provided corrective feedback regarding proper implementation; (c) showed the teacher a graph of teacher treatment fidelity; (d) provided the teacher encouragement to continue implementation; (e) reminded the teacher that the researcher would be back the next day.

Procedural fidelity was collected during self-monitoring phases similar to the performance feedback sessions (see Appendix B). The researcher observed the self-monitoring session between the consultant and teacher and noted if the steps were performed or not. The self-monitoring procedural fidelity steps were (a) compare the consultant treatment integrity form to the teacher self-monitoring form; (b) list all steps in which there is agreement between the consultant and the teacher; (c) discuss any

disagreements and come to a consensus on who is accurate; (d) provide corrective feedback regarding disagreements (if any); and (d) provide encouragement to continue self-monitoring. Procedural fidelity was calculated as the number of procedures that were completed, divided by the total number of procedures possible and multiplied by 100 to obtain a percentage.

Social Validity Measures

Social validity was assessed through questionnaires provided to the teacher before and after the study was concluded (see Appendix C). Teachers rated the GBG and self-monitoring procedures separately on a scale of 1 to 5 with lower numbers indicating less social validity and higher numbers indicating higher social validity. Teachers rated items relevant to (a) overall satisfaction (b) the acceptability of the intervention and time required; and (c) their satisfaction with overall positive effects.

Experimental Design

A multiple baseline across participants design was used to examine the effects of performance feedback and self-monitoring on treatment integrity. Multiple baseline designs are demonstration designs in which the independent variable is sequentially introduced across baselines (Baer, Wolf, Risley, 1968; Kennedy, 2005). Replication of a functional relation occurs across multiple tiers. Because lower tiers remain in the baseline condition while the independent variable is applied to higher tiers, history and maturation effects can be detected and controlled. This design is most appropriately used for behaviors that are non-reversible or where reversing the behavior is not desirable (Kennedy, 2005).

Procedures

Baseline. The baseline condition occurred during language arts period for 10 minutes. Each teacher conducted lessons using the McMillan/McGraw-Hill *Treasures Series* curriculum. Supplemental materials included teacher made worksheets. All teachers used a classroom management system based on a card system. Each student had a series of colored cards kept in a pouch at the front of the classroom visible to all students. If inappropriate behavior occurred, the teacher would tell the student to pull a card. Students started out on a green card but as each card is pulled, the color changed indicating the student's level. If a student progressed from green to red, the student was sent to the office for disciplinary action. Teacher 2 used a token economy system in addition to the card system. These tokens were kept in a pouch around her waist for distribution. During baseline, the GBG was not implemented.

Training on the Good Behavior Game. Training occurred after school in the teacher's classroom in a didactic training environment between the consultant and teacher and lasted approximately 60 min (range 60 – 90 min). No students were present. Teachers were given content related to the GBG through handouts and oral presentation. Specifically, the first author as the consultant described the rationale and research behind the GBG and each of the 14 steps that would be monitored for treatment fidelity. The consultant simulated the game in front of the classroom for the participants without students present. The consultant described and modeled the procedures for handling rule violations as if students were present. All participants then watched a video of the GBG being played in a typical classroom by a trained teacher. Participants watched the video a second time, rating the teacher in the video using the treatment fidelity checklist. The consultant and participants discussed steps the teacher in the video implemented

effectively and steps either not implemented or implemented with lower quality than desired by reviewing the fidelity checklist and pertinent sections of the video. The exception was for Teacher 4 who did not have the opportunity to view the video clip. Finally, the participants role-played the GBG without students present and with consultant feedback for 10 min until the participant performed each step of the GBG with 100% fidelity. The consultant told the participants after approximately a week of performance feedback they would be monitoring themselves using a simplified version of the treatment fidelity checklist.

Good Behavior Game with performance feedback. Performance feedback was provided on the initial day of the GBG phase (after training) and continued daily for at least five data points when a phase change occurred. Teachers did not self-monitor during this phase. Teachers were told they could continue to use the current classroom management strategies (i.e., “pull a card”) during this phase although no teachers were observed using this system during the observational time of this phase. Performance feedback consisted of a 3-5 min meeting between the consultant and the teacher immediately following implementation of the GBG during the language arts period. During performance feedback, the consultant showed the teacher the scored treatment fidelity sheet and orally (a) reviewed steps performed correctly and provide specific praise for at least one step (e.g., “You performed all of the beginning and ending steps correctly. I really liked the way you asked students to identify their team membership by giving the animal sounds associated with their team names.”); (b) described any steps that were missing or incorrect and provided corrective feedback regarding proper implementation (e.g., “Although you responded immediately and to every rule violation, you forgot to praise the other teams who did not receive the rule violation. Remember to

say something like ‘Thank you Lions and Tigers for raising your hand to speak’.”); (c) showed the teacher a graph of her teacher treatment fidelity data; (d) asked the teacher to continue implementation; and (e) told the teacher that the consultant would back the next day. A sample script of performance feedback sessions is provided in Appendix A.

Good Behavior Game with self-monitoring. Once treatment integrity data were stable and above 80% treatment fidelity for five consecutive sessions, a phase change occurred. Performance feedback was no longer provided and the teacher self-monitored treatment fidelity using a simplified treatment fidelity form previously used by the consultant during the performance feedback phase (Appendix A). The consultant explained the use of the treatment fidelity form which was familiar from the performance feedback sessions. The participant was told to look at each step and check *Yes* or *No* regarding whether the step was performed. For steps in which there could be multiple opportunities (e.g., *responding immediately to every rule violation*), the participants were told to check *Yes* if they performed that step more often than not (i.e., more than half of the time). Additionally, participants were told to make a note in the comments section if there were areas they felt needed improvement (e.g., missed one instance of *responding immediately to every rule violation*). As was the case during the performance feedback phase, participants were told they could continue to use their current classroom management strategy although this did not occur.

For the first three sessions of this phase, the consultant and teacher concurrently collected treatment fidelity data and compared results to determine interobserver agreement (IOA). Once the teacher and consultant agreed on 100% of the items regardless of the overall percentage of treatment fidelity, consultant support was systematically faded. In other words, if a teacher missed a step the overall percentage of

treatment fidelity would be 93% of steps implemented. The teacher and consultant were required to agree that treatment fidelity was 93% for steps implemented. After the consultant and teacher reached 100% agreement for three consecutive days, the teacher continued to implement the GBG and self-monitor but the consultant did not return to the classroom. The consultant told the teachers she would return at unannounced times to conduct fidelity agreement checks. After two days, the consultant returned to compare treatment fidelity data with the teacher. If treatment fidelity agreement between the consultant and teacher was above 80%, the consultant did not return to the classroom for three more days at which time another fidelity agreement check was performed between the teacher and consultant. If agreement on treatment fidelity remained above 80%, the consultant conducted a final fidelity agreement check after three additional days and prior to initiation of the maintenance phase. Teachers were asked to leave completed self-monitoring sheets in a pre-determined area for the consultant to collect during random times over the course of the self-monitoring phase.

Maintenance. During the maintenance phase, teachers were told that the consultant would no longer be observing implementation of the GBG and would not be collecting self-monitoring sheets. Teachers were provided additional self-monitoring sheets and told they could continue using the sheets if they so desired. At the conclusion of the maintenance phase, the consultant returned each classroom to ask the teachers if they had continued to play the game and if so, whether or not they continued to use the self-monitoring sheets.

CHAPTER III

RESULTS

Formative Evaluation

Observational data were collected across all phases and participants. The main dependent variable was the percentage of treatment fidelity steps observed. The dependent measure was displayed in a time series graph (Figure 1). Visual analysis of the percentage of treatment fidelity was used to make all formative decisions regarding stability, level, and trend of data and phase changes.

Interobserver agreement and procedural fidelity. Interobserver agreement and procedural fidelity data were collected during all phases for 41% of observations. Interobserver agreement was 100% for baseline across Teachers 1 through 4 and 100% across phases for Teachers 1 and 4. The IOA during performance feedback phase for Teacher 2 averaged 98% (range 93-100) and 98.6% (range 93-100) for Teacher 3. During the teacher self-monitoring phase, IOA remained at 100% across teachers. Procedural fidelity across phases remained at 100%, and, therefore, booster training sessions were not required.

Baseline. Baseline observations occurred for 10 min during each teacher's language art period between 10:00am – 11:00am daily. The timing of the observations was constant within each classroom. For example, Teacher 1 was observed approximately between 10:00-10:15, Teacher 2 between 10:15-10:30, Teacher 3 between 10:30 and 10:45 and Teacher 4 between 10:45 and 11:00. Treatment fidelity of the GBG

was measured and remained zero across participants for each baseline observation (Figure 1).

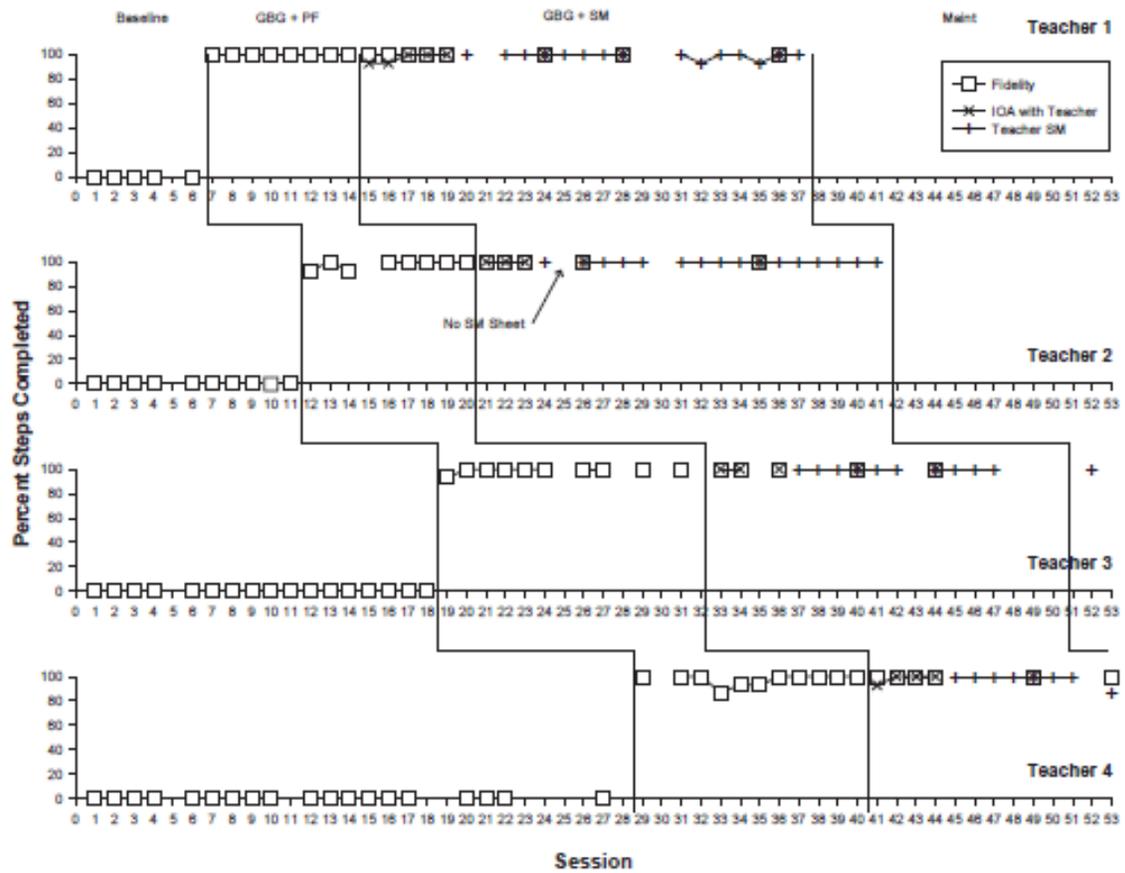


Figure 1. Performance Feedback and Teacher Self-Monitoring on Treatment Fidelity of the GBG

Agreement was also collected for 41% of the observation sessions and was 100% across participants. Four data points were collected prior to school break. One additional observation occurred for Teacher 1 after break prior to initiation of the first phase change.

Good Behavior Game with performance feedback. Introduction of the independent variable occurred in a step-wise fashion across all participants. Decisions regarding phase changes were made based on the stability of the dependent variable and IOA at 80% or above. Prior to a phase change from baseline to GBG plus performance feedback, teachers received didactic training on the GBG as described previously. An immediate shift in treatment fidelity occurred across participants once the independent variable was introduced. Teacher 1 obtained 100% treatment fidelity quantity immediately and remained stable across the phase with a mean quality fidelity rating of 99% (range 97 – 100%). Teacher 2 also obtained immediate high treatment fidelity quantity with slight variability in the first three data points fluctuating between 93 – 100%. After the third observation session, Teacher 2 remained at 100% treatment fidelity quantity for the duration of the phase. The mean treatment fidelity quality rating for Teacher 2 was 96% (range 91 – 100%). Similarly, Teacher 3 obtained 93% treatment fidelity but then remained at 100% for the duration of the phase with a mean treatment fidelity quality of 98% (range 91-100). Teacher 4 had the greatest variation in treatment fidelity quantity and quality across teachers. Teacher 4 obtained 100% treatment fidelity quantity initially but then dropped to 85% (missing two steps) for one observation and increased to 93% for two additional sessions before returning to 100% treatment fidelity for the remainder of the phase. The mean treatment fidelity quality rating for Teacher 4 was 97% (87 – 100%). Agreement during GBG plus performance feedback remained

above 90% for fidelity quantity and 86% for fidelity quality. Procedural fidelity remained at 100% across the phase.

Good Behavior Game with self-monitoring. Once participants obtained 100% fidelity for at least five sessions, a phase change occurred to self-monitoring. In general, fidelity quantity and quality remained high and stable across teachers. Teachers were required to obtain 100% reliability on treatment fidelity quantity with the consultant for three consecutive sessions before self-monitoring on their own. Teachers 2 and 3 achieved agreement immediately for three consecutive sessions whereas Teachers 1 and 4 scored themselves lower on treatment fidelity compared with the consultant. Once agreement was obtained for three consecutive sessions, observations did not occur for at least three sessions. Fidelity quantity remained at 100% across teachers during fidelity checks by the consultant. Similarly Teacher 1, 3, and 4 all obtained 100% treatment fidelity quality across the phase with Teacher 2 obtaining $M = 99%$ (range 91 – 100). Teacher self-monitoring sheets indicated 100% fidelity by Teachers 2 and 3 across the self-monitoring phase. Teachers 1 and 4 reported 93% fidelity for two sessions and one session respectively. Procedural fidelity and IOA remained at 100% across the phase and across participants.

Maintenance. After three fidelity checks by the consultant a phase change to maintenance occurred. Observations did not occur and the consultant stopped in at the conclusion of the school year. Teachers 1 and 2 reported continued use of the GBG although they did not use the self-monitoring sheets. Teacher 3 used the self-monitoring sheet once and reported she continued playing the game until the last few days of school when other demands prevented her from playing it. Due to the conclusion of the school year, Teacher 4 did not reach the maintenance phase.

Summative Evaluation

An analysis of the data across participants was conducted to determine if a functional relation exists between treatment fidelity and performance feedback and performance feedback and self-monitoring. First, IOA and procedural fidelity were examined to determine if both were above 80% across participants and phases. Procedural fidelity remained at 100% across participants and phases and quantity IOA remained at or above 93% across participants and phases while the quality IOA remained above 86% ($M = .98$) across participants and phases. This indicates no threats to instrumentation were evident and procedural fidelity was adequate to continue further analysis of the data.

Visual analysis was conducted to examine shifts in data both within condition and between conditions. There were no within condition shifts in data during baseline for all participants. Baseline levels of treatment fidelity of the GBG remained at zero for all participants throughout the study. An examination of the GBG plus performance-feedback phase across participants shows stability across participants with the greatest shift in Teacher 4 between 87-100% ($M = .98$). A similar pattern occurred during GBG and self-monitoring phases. Treatment fidelity remained high and stable once teachers began self-monitoring and consultant support was faded.

Between condition shifts were analyzed to determine if they were greater than any within condition shifts. Data shifted immediately and in the therapeutic direction across all participants when the GBG and performance feedback was applied with no shifts occurring for participants when the dependent variable was not applied. Altogether this evaluation supports the conclusion that a functional relation exists between treatment

fidelity of the GBG and use of performance feedback to achieve high fidelity and sustaining high fidelity using teacher self-monitoring.

Social Validity

Prior to initiation of the study, ratings of the GBG were generally positive across each category ($M = 3.88$; $SD = 0.33$) with the lowest ratings for *Fit/Ease* ($M = 3.58$) and highest ratings for *Time/Burden* ($M = 4.25$). Across teachers, the overall mean rating ranged from 3.71 to 4.08 ($M = 3.88$; $SD = 0.15$) with the lowest rating coming from Teacher 4 and the highest rating from Teacher 1. A similar positive pattern emerged for social validity ratings of self-monitoring across categories ($M = 3.98$; $SD = 0.34$) with the lowest rating for *Fit/Ease* ($M = 3.63$) and the highest rating for *Overall support* ($M = 4.31$). Teacher 2 rated self-monitoring the lowest ($M = 3.67$) and Teacher 4 rated self-monitoring the highest ($M = 4.33$).

After the conclusion of the study, participants completed the same social validity forms. In general, teachers rated the GBG higher after the study ($M = 4.03$; $SD = 0.36$) with the lowest ratings for *Produces positive effects* ($M = 3.6$) and the highest ratings for *Overall support* ($M = 4.5$). Across teachers, the overall mean rating ranged from 2.85 – 4.31 ($M = 4$) with the lowest rating from Teacher 3 and the highest rating from Teacher 2. The post social validity ratings of self-monitoring across categories in general increased slightly ($M = 4$; $SD = 0.33$) with the lowest ratings for *Time/Burden* ($M = 3.75$) and the highest rating for *Overall support* ($M = 4.38$). As was the case for post measures of the GBG, Teacher 3 rated self-monitoring the lowest ($M = 3.33$) while Teacher 4 rated self-monitoring the highest $M = 4.33$). Table 1 provides the means, standard deviations, and ranges of scores pre- and post-test for the Good Behavior Game and self-monitoring. Tables 2 through 5 provide a breakdown of pre- and post-test social validity ratings of the

Good Behavior Game and self-monitoring based on the number of teachers' ratings for each question.

Table 1
 Pre- and Post-test Social Validity Ratings of the Good Behavior Game and Self-Monitoring

	Good Behavior Game				Self-Monitoring			
	<i>M</i>	SD	Range		<i>M</i>	SD	Range	
			Low	High			Low	High
Pre	3.88	0.33	3.71	4.08	3.98	0.34	3.67	4.33
Post	4.03	0.36	3.60	4.50	4.00	0.33	3.75	4.38

Table 2

Pre-test Social Validity Ratings of the Good Behavior Game Based on the Number of Teachers' Ratings for each Question.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
		1	2	3	4	5
Overall Support	1. I received adequate training on how to use the Good Behavior Game in my classroom				3	1
	2. I understood how to use the Good Behavior Game.				4	
	3. I enjoyed implementing the Good Behavior Game.				4	
	4. I plan to use the Good Behavior Game in my classroom in the future.				4	
Fit/Ease	5. I found it easy to use the Good Behavior Game in my classroom.				4	
	6. I do not know of any other intervention that would be a better fit for my students.		1	2	1	
	7. The Good Behavior Game was a good fit for students in my classroom.			1	3	
Time/Burden	8. Adding the Good Behavior Game did not interfere with academic instruction and routines in my classroom.				3	1
	9. Using the Good Behavior Game did not take up too much of my time.				3	1
Produces positive effects	10. After starting the Good Behavior Game in my classroom, I was able to see immediate changes in my students' behavior.			1	3	
	11. The addition of the Good Behavior Game has improved academics in my classroom.		1		3	
	12. The addition of the Good Behavior Game has improved behavior in my classroom.		1		3	
	13. The addition of the Good Behavior Game has improved the atmosphere in my classroom.			1	3	

Table 3

Post-test Social Validity Ratings of the Good Behavior Game Based on the Number of Teachers' Ratings for each Question.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
		1	2	3	4	5
Overall Support	1. I received adequate training on how to use the Good Behavior Game in my classroom				1	3
	2. I understood how to use the Good Behavior Game.				1	3
	3. I enjoyed implementing the Good Behavior Game.			1	1	2
	4. I plan to use the Good Behavior Game in my classroom in the future.			1	1	2
Fit/Ease	5. I found it easy to use the Good Behavior Game in my classroom.				1	3
	6. I do not know of any other intervention that would be a better fit for my students.		1	1	2	
	7. The Good Behavior Game was a good fit for students in my classroom.			1	2	1
Time/Burden	8. Adding the Good Behavior Game did not interfere with academic instruction and routines in my classroom.		1		1	2
	9. Using the Good Behavior Game did not take up too much of my time.		1		1	2
Produces positive effects	10. After starting the Good Behavior Game in my classroom, I was able to see immediate changes in my students' behavior.			1	3	
	11. The addition of the Good Behavior Game has improved academics in my classroom.			1	2	1
	12. The addition of the Good Behavior Game has improved behavior in my classroom.		1		2	1
	13. The addition of the Good Behavior Game has improved the atmosphere in my classroom.		1		2	1

Table 4
Pre-test Social Validity Ratings of Self-Monitoring Based on the Number of Teachers' Ratings for each Question.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
		1	2	3	4	5
Overall Support	1. I received adequate training on how to my use of the Good Behavior Game in my classroom				2	2
	2. I understood how to self-monitor my use of the Good Behavior Game.				3	1
	3. I enjoyed self-monitoring my use of the Good Behavior Game.				3	1
	4. I plan to use self-monitoring in my classroom in the future.				3	1
Fit/Ease	5. I found it easy to self-monitor in my classroom.				2	2
	6. I do not know of any other intervention that would be a better fit for me.			3	1	
Time/Burden	7. Adding self-monitoring did not interfere with academic instruction and routines in my classroom.			1	3	
	8. Using self-monitoring of the Good Behavior Game did not take up too much of my time.				3	1

Table 5
Post-test Social Validity Ratings of Self-Monitoring Based on the Number of Teachers' Ratings for each Question.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
		1	2	3	4	5
Overall Support	1. I received adequate training on how to my use of the Good Behavior Game in my classroom				1	3
	2. I understood how to self-monitor my use of the Good Behavior Game.				1	3
	3. I enjoyed self-monitoring my use of the Good Behavior Game.			1	2	1
	4. I plan to use self-monitoring in my classroom in the future.			1	2	1
Fit/Ease	5. I found it easy to self-monitor in my classroom.				2	2
	6. I do not know of any other intervention that would be a better fit for me.		1	1	2	
Time/Burden	7. Adding self-monitoring did not interfere with academic instruction and routines in my classroom.		1		2	1
	8. Using self-monitoring of the Good Behavior Game did not take up too much of my time.			2	1	1

CHAPTER IV

DISCUSSION

Improving student outcomes is enhanced by the successful adoption of evidence-based practices by classroom teachers (The Evidence-Based Intervention Work Group, 2005). Successful adoption requires teachers to implement evidence-based practices as intended, or with high fidelity, to achieve student outcomes documented from previous research on such evidence-based practices (Dusenbury et al., 2003). Supporting teachers' use of EPBs can be facilitated by performance feedback; however fidelity of implementation tends to diminish once feedback is withdrawn (Noell et al., 1997; Witt et al., 1997). Given prior research it was hypothesized that self-monitoring could be a viable method to fade performance feedback support while still maintaining high levels of treatment fidelity (Kissel et al., 1983; Petscher & Bailey, 2006). The purpose of this study was to examine if treatment fidelity could be maintained using teacher self-monitoring once researcher support was withdrawn.

Didactic training between the consultant and teachers on the Good Behavior Game, an evidence-based classroom management practice, produced an immediate increase in treatment fidelity above 90% across teachers indicating its effectiveness at producing acquisition of GBG implementation skills. Performance feedback was used to continue high levels of treatment fidelity and fluency of GBG implementation. Across teachers, treatment fidelity above 80% was achieved relatively quickly with performance feedback. Three of the four teachers reached and maintained treatment fidelity above

90% within three sessions. The other teacher had three sessions in which treatment fidelity dropped but remained above 80% and then subsequently returned to 100% steps implemented. Once performance feedback was terminated and self-monitoring initiated, teachers maintained a consistent level of fidelity typically around 100% steps implemented. Fidelity checks occurred after several days of teachers' continued use of self-monitoring of the GBG without performance feedback. These checks also indicated that despite the lack of researcher feedback and support, teachers were able to maintain high levels of fidelity. The results from this study indicate self-monitoring may be a viable method to fade support.

Interestingly, the differences between teachers and researchers in agreement of steps implemented occurred due to teachers' lower ratings of themselves rather than rating themselves higher than actual levels of fidelity. For the two teachers that rated themselves lower than the consultant did, this typically occurred for procedures to handling disruptive behaviors. The two steps teacher rated themselves lower were (1) responding immediately to rule violation and (2) responding consistently to every rule violation. Teachers reported concerns that they may have missed instances of rule violations when their backs were turned or they were distracted. In such cases, they would rate themselves as not having done these steps, rather than indicating it was performed more often than not although the quality could be improved. This may indicate a higher level of diligence by adults in rating their own behavior.

Measures of social validity in general were positive and scores were higher post study. However, there were differences in ratings pre- and post within teachers, namely Teacher 3. The differences in social validity ratings of self-monitoring for Teacher 3 remained roughly the same however there were two notable categories where Teacher 3

rated the GBG lower after the conclusion of the study. Teacher 3 rated GBG lower for *Time/Burden* (Pre $M = 4$ compared to Post $M = 2$) and *Produces positive effects* (Pre $M = 4$ compared to Post $M = 3$). At least two explanations are possible. First, Teacher 3 reported to the researcher she was reluctant to participate in the study because she did not like being watched but chose to participate in the hopes that it would lessen her anxiety about being observed. Receiving performance feedback from an observer (i.e., the consultant) may have contributed to her lower rating of the GBG. Second, the study continued until the end of the school year when multiple demands are placed on teachers to complete lessons. Teacher 3 reported not conducting the GBG due to burdens to finish student testing. These factors may have influenced Teacher 3's social validity ratings.

All teachers reported their students enjoyed the game and it reduced disruptive student behavior while the game was being played. However, teachers also reported dismay that the same results did not happen across the school day when the game was not being played. This indicates the skills the teachers learned by implementing the game did not generalize to other times of the day. Teachers individually told the consultant students "can behave when the game is being played but the minute the game stops, they return to misbehaving." This may indicate an inability for the teachers to recognize how their own behavior plays an important role in student behavior and programming for teacher generalization of behaviors needs to occur so that positive effects can be obtained across the school day.

The results from this study indicate support for teachers is highly desirable and necessary to achieve high fidelity of implementation of EBP like the GBG. Teachers highly rated the overall support they received for both the GBG and self-monitoring prior to the initiation of the study and after. Additionally, teacher reported that the performance

feedback they received from the researcher prior to self-monitoring was necessary for accurate implementation of the GBG. Results from this study indicate that combining daily performance feedback with teacher self-monitoring may be an effective approach to sustain high fidelity of teacher implementation of evidence-based classroom management practices.

Limitations and Future Research

One limitation of this study is that the maintenance phase was much shorter than desired due the end of the school year. Ideally, maintenance periods for teacher implementation of evidence-based practices would occur for roughly eight weeks or longer. In this study, maintenance was three to four weeks for the first two teachers, approximately two weeks for Teacher 3, and did not occur for Teacher 4. Several questions remain unanswered. For example, was teachers' continued use of self-monitoring sheets necessary to maintain fidelity over long periods of time? During the maintenance phase, teachers did not report continuing use of the self-monitoring sheets. Anecdotal reports indicate the teachers chose not to continue filling out the sheets because it was "one less thing to worry about" or that they "self-monitored in their own head" rather than using a sheet. It is not clear if this is an indication of the usefulness or lack of usefulness of the self-monitoring sheets or if teachers became proficient enough that self-monitoring occurred seamlessly or without thought. In other words, teachers began supervising their own practice (Kilbourn, 1991). The end of the school year may have been an additional factor in teachers' decision not to use the self-monitoring forms. If the maintenance phase occurred for a longer period of time, it may be possible to determine this with fidelity checks temporally distant from performance feedback. Reductions in fidelity after teachers discontinued using the self-monitoring sheets may

indicate they served as a fidelity prompt whereas continued high levels of fidelity may be more of an indication that teacher fluency and proficiency with implementation of the GBG. Researchers should consider replicating this study with an extended maintenance phase to examine if the use of teacher self-monitoring after a brief period of performance feedback can maintain treatment fidelity for longer periods of time, ideally over a school year. Moreover, research should be conducted on the use of teacher self-monitoring after performance feedback with other evidence-based practices that are more or less complex than the GBG.

Another limitation of this study is the lack of direct measures of student on-task or disruptive behavior. These measures were not collected because reductions in disruptive student behavior and increases in student time on-task have been reported in the literature when the GBG was implemented in classrooms. These results were therefore expected in the current study. The addition of student measures of disruptive behavior or time on-task may provide additional data to examine if fluctuations in inappropriate student behavior co-occur for absent or lower ratings of specific fidelity steps. For example, increases in disruptive behavior may occur when the teacher does not praise the other teams after a rule violation while playing GBG indicating this particular step may necessary for greater results. On the other hand, if no changes in student disruptive behavior occur despite the fact that the teacher did not praise other teams after a rule violation, it may indicate a lack of importance of this particular step to rates of disruptive student behavior. In the future, researchers should consider collecting direct measures of student behavior to examine if variations in student behavior coincide with absent or lower ratings of specific steps of treatment fidelity.

Conclusion

The purpose of the current study was to determine if the use of teacher self-monitoring after a brief period of performance feedback could be an efficient means for teachers to maintain high levels of an evidence-based classroom management practice. Results from this study indicate teachers' use of self-monitoring did in fact maintain high levels of treatment fidelity obtained through didactic training and subsequent performance feedback. As research in the area of teacher support of implementation of evidence-based practice continues, researchers should identify optimal support required to effectively implement said practices. As researchers, when we recommend a practice as evidence-based, it is incumbent upon us to also provide research on the most efficient way support teacher implementation of those practices. As school systems make decisions about the adoption of EBP knowledge of the supports required to implement those EBP with fidelity will not only aid in the decision making process, but in ensuring that proper supports are in place for successful adoption and maintenance.

Appendix A
Materials and Treatment Fidelity Forms

Good Behavior Game Chart

Goal: ___ points or less to win!

Teams	M	T	W	Th	F	Weekly Winner
Behaviors: • • • •						
Place a  next to the winning team(s)						

GOOD BEHAVIOR GAME: TREATMENT INTEGRITY FORM

Teacher:	Observer:		
School:	Date:		
Start time:	End time:	Total time:	min/day
How would you categorize this observation of implementation?	___(0) Observation took place, teacher not implementing	___(2) No observation, teacher cancel	
	___(1) Observation took place with Teacher Implementing	___(3) No observation, consultant was unavailable	
		___(4) No observation, school on break or other school related interruption	
		___(5) Consultant available but teacher not in classroom	

	Observed		If not Observed, Other Evidence		Quality Rating: <small>Based on observation or teacher report, rate the degree to which the teacher implemented of each component</small>				
	Check yes or no		Describe:		Not Implemented	Minimal Fidelity	Half/Partial Fidelity	Good Fidelity	Full Fidelity
	Yes	No	Yes	No					
Getting Ready to Begin									
1. Announce game before beginning.					1	2	3	4	5
2. Refer to teams before beginning.					1	2	3	4	5
3. Refer to the classroom rules.					1	2	3	4	5
4. Refer to requirements to win.					1	2	3	4	5
5. Refer to the rule violation process.					1	2	3	4	5
6. Indicate the start of the game.					1	2	3	4	5
Handle disruptive behaviors by:									
7. a. Responding immediately to rule violation					1	2	3	4	5
b. Responding consistently to every rule violation					1	2	3	4	5
8. Responding with a typical instructional voice					1	2	3	4	5
9. Identifying the child or team					1	2	3	4	5
10. Praising the other children or team(s)					1	2	3	4	5
At the end of the game:									
11. Review scores.					1	2	3	4	5
12. Hand out prizes or deliver other reinforcers					1	2	3	4	5
Overall									
13. Game is played with a clear beginning and end (timer or discrete instructional time used)					1	2	3	4	5
Totals									
Positive Comments:									
Target(s) for Improvement (if any):									

Note: If there was no opportunity to observe the behavior (e.g., no disruptive behavior to respond to) write NA in the column for observed "yes" and do not circle a quality rating.

**GOOD BEHAVIOR GAME: TREATMENT INTEGRITY
SELF MONITORING FORM**

Teacher:	School:	Date:
Start time:	End time:	Total time: min.

	Did I?		Comments
	Yes	No	
Getting ready to begin			
1. Announce game before beginning.			
2. Refer to teams before beginning.			
3. Refer to the classroom rules.			
4. Refer to requirements to win.			
5. Refer to the rule violation process.			
6. Indicate the start of the game.			
Handle Disruptive Behaviors By:			
7. a. Responding immediately to rule violation			
b. Responding consistently to every rule violation			
8. Responding with a typical instructional voice			
9. Identifying the child or team			
10. Praising the other children or team(s)			
At the end of the game:			
11. Review scores.			
12. Hand out prizes or deliver other reinforcers			
Overall			
13. Game is played with a clear beginning and end (timer or discrete instructional time used)			
How did I do? Total Treatment Fidelity	/14	X 100 =	%
Goals for tomorrow:			

GOOD BEHAVIOR GAME: TREATMENT INTEGRITY FORM

Teacher: Mrs. Smith

Observer: Ms. Consultant

School: Hermitage Elementary

Date: 2/11/10

Start time: 9:05 am

End time: 9:25am

Total time: 20 min/day

How would you categorize this observation of implementation?

(0) Observation took place, teacher not implementing
 (1) Observation took place with Teacher Implementing

(2) No observation, teacher cancel
 (3) No observation, consultant was unavailable
 (4) No observation, school on break or other school related interruption
 (5) Consultant available but teacher not in classroom

	Observed		If not Observed, Other Evidence		Quality Rating:					
	Check yes or no		Describe:		Based on observation or teacher report, rate the degree to which the teacher implemented of each component					
	Yes	No	Yes	No	Not Implemented	Minimal Fidelity	Half/Partial Fidelity	Good Fidelity	Full Fidelity	
Getting Ready to Begin										
1. Announce game before beginning.	X				1	2	3	4	5	
2. Refer to teams before beginning.		X			1	2	3	4	5	
3. Refer to the classroom rules.	X				1	2	3	4	5	
4. Refer to requirements to win.	X				1	2	3	4	5	
5. Refer to the rule violation process.	X				1	2	3	4	5	
6. Indicate the start of the game.	X				1	2	3	4	5	
Handle disruptive behaviors by:										
7. Responding immediately to rule violation	X				1	2	3	4	5	
8. Responding consistently to every rule violation	X				1	2	3	4	5	
9. Responding with a typical instructional voice	X				1	2	3	4	5	
10. Identifying the child or team	X				1	2	3	4	5	
11. Praising the other children or team(s)	X				1	2	3	4	5	
At the end of the game:										
12. Review scores.	X				1	2	3	4	5	
13. Hand out prizes or deliver other reinforcers	X				1	2	3	4	5	
Overall										
14. Game is played with a clear beginning and end (timer or discrete instructional time used)	X				1	2	3	4	5	
Totals	13	1								
<p>Positive Comments: Great job using the timer and giving behavior specific praise. Seemed more confident.</p> <p>Target(s) for Improvement (if any): Remember to refer to teams and praise other children when responding to disruptive behavior.</p>										

Note: If there was no opportunity to observe the behavior (e.g., no disruptive behavior to respond to) write NA in the column for observed "yes" and do not circle a quality rating.

Sample Consultant Script for Performance Feedback

- (1) List each step that was performed correctly and provide praise for at least one.
- (2) Describe any steps that were missing or incorrect and provide corrective feedback regarding proper implementation.
- (3) Show the teacher a graph of student disruptive behavior and teacher treatment fidelity.
- (4) Provide encouragement to continue implementation.
- (5) Remind the teacher that the consultant will be back the next day.

“Hello Mrs. Smith. That was another really great Good Behavior Game session. You completed all of the steps except for one and I really like your use of behavior specific praise. In terms of getting ready to begin, you correctly announced the game before beginning and referred to the classroom rules, requirements to win, and the rule violation process. I also liked your use of the timer to indicate the start and end of the game. That seems to be working much better for you.

You also did a much better job handling disruptive behaviors from last time. I noticed you responded immediately to rule violations and in general, you were consistent with catching each violation and using an instructional voice. You identified the team or student that was disruptive and praised the other students or team.

At the end of the game, you reviewed scores and gave the teams’ free time to play checkers which is what they had selected. Using the timer really made a difference making a clear definition of the beginning and end of the game.

The only corrective feedback I have for you relates to referring to the teams before the beginning of the game. Remember when you announce that the game is going to begin, you should remind the students what team they are on so there is no question. List each student’s name that is on Team A and then list each student’s name that is on Team B. In terms of the quality ratings, I noticed that you forgot to mention that both teams could win the game when you went over the requirements. I also noticed that there was one disruptive behavior that you did not respond to, when Sam made that noise with his mouth.

Here is the graph of your treatment fidelity scores. You’ve consistently have increased your percentage since we started and I expect that you will be at 100% tomorrow given what I saw today. We can also see a continued decrease in student disruptive behavior as you can see by the graph. It started out pretty high but quickly dropped and is continuing to decrease. It seems like the game is really working. You’re doing a great job and I want to encourage you to keep playing the game and working on your fidelity. Do you have any questions? Ok, I’ll be back tomorrow to observe and give you feedback. Thank you.”

Appendix B
Procedural Fidelity Forms

PERFORMANCE FEEDBACK: PROCEDURAL FIDELITY FORM

Consultants:	Researcher:	Date:
Teacher:	School:	
Start time:	End time:	Total time: min.

	Yes	No	Comments
Performance Feedback Steps			
1. List each step that was performed correctly and provide praise for at least one.			
2. Describe any steps that were missing or incorrect and provide corrective feedback regarding proper implementation.			
3. Show the teacher a graph of student disruptive behavior and teacher treatment fidelity			
4. Provide encouragement to continue implementation.			
5. Remind the teacher that the consultant will be back the next day.			
Total:	/5		

SELF-MONITORING: PROCEDURAL FIDELITY FORM

Consultants:	Researcher:	Date:
Teacher:	School:	
Start time:	End time:	Total time: min.

	Yes	No	Comments
Self-Monitoring Steps			
1. Compare the consultant treatment integrity form to the teacher's self-monitoring form.			
2. List all steps in which there is agreement between the teacher and consultant.			
3. Discuss any disagreements and come to a consensus on who is accurate.			
4. Provide corrective feedback regarding disagreements (if any).			
5. Provide encouragement to continue self monitoring.			
Total:	/5		

Appendix C
Social Validity Forms

**Social Validity Rating Form: Performance Feedback
Good Behavior Game**

Teacher: _____
School: _____

Date: _____
District: _____

Circle the number reflecting your level of agreement for each statement.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Overall support	1. I received adequate training on how to use of the Good Behavior Game (GBG).	1	2	3	4	5
	2. I understood how performance feedback would enhance my implementation of the GBG.	1	2	3	4	5
	3. I enjoyed receiving performance feedback on my use of the GBG.	1	2	3	4	5
	4. I plan to seek out performance feedback in my classroom in the future.	1	2	3	4	5
Fit/Ease	5. I found it easy to receive performance feedback.	1	2	3	4	5
	6. I do not know of any other interventions that would be a better fit for me.	1	2	3	4	5
Time/Burden	7. Adding performance did not interfere with academic instruction and routines in my classroom.	1	2	3	4	5
	8. Receiving performance feedback of the GBG did not take up too much of my time.	1	2	3	4	5
Produces positive effects	9. After receiving performance feedback, I was able to see immediate changes in my use of the GBG.	1	2	3	4	5
	10. The addition of performance feedback has improved my use of the GBG.	1	2	3	4	5

Comments:

**Social Validity Rating Form: Self-Monitoring
Good Behavior Game**

Teacher: _____
School: _____

Date: _____
District: _____

Circle the number reflecting your level of agreement for each statement.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Overall support	1. I received adequate training on how to self-monitor my use of the Good Behavior Game.	1	2	3	4	5
	2. I understood how to self-monitor my use of the Good Behavior Game.	1	2	3	4	5
	3. I enjoyed self-monitoring my use of the Good Behavior Game.	1	2	3	4	5
	4. I plan to use self-monitoring in my classroom in the future.	1	2	3	4	5
Fit/Ease	5. I found it easy to use self-monitoring in my classroom.	1	2	3	4	5
Time/Burden	6. Adding self-monitoring did not interfere with academic instruction and routines in my classroom.	1	2	3	4	5
	7. Using self-monitoring of the Good Behavior Game did not take up too much of my time.	1	2	3	4	5

Comments:

REFERENCES

- Alberto, P. A., & Troutman, A. C. (2009). *Applied behavior analysis for teachers (8th ed.)*. Upper Saddle River, NJ: Merrill/Pearson.
- Allinder, R., & BeckBest, M. (1995). Differential effects of two approaches to supporting teachers' use of curriculum-based measurement. *School Psychology Review, 24*, 287-298.
- Allinder, R., Bollins, R., Oats, R., & Gagnon, W. (2000). Effects of teacher self-monitoring on implementation of curriculum-based measurement and mathematics computation achievement of students with disabilities. *Remedial and Special Education, 21*, 219-226.
- Alvero, A., Bucklin, B., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings. *Journal of Organizational Behavior Management, 21*, 3-29.
- American Federation of Teachers (2000). Building on the best, learning from what works: Five promising discipline and violence prevention programs. Washington, DC: Author.
- Angrist, J., & Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review, 27*, 483-503. doi: 10.1016/j.econedurev.2007.03.002
- Baer, D., Wolf, M., & Risley, T. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97.

- Barrish, H., Saunders, M., & Wolf, M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*, 119-124.
- Belfiore, P. J., & Browder, D. M. (1992). The effects of self-monitoring on teachers' data-based decisions and on the progress of adults with severe mental retardation. *Education and Training in Mental Retardation, 27*, 60-67.
- Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education, 56*, 205-213. doi: 10.1177/0022487105275904
- Bingham, M., Spooner, F., & Browder, D. (2007). Training paraeducators to promote the use of augmentative and alternative communication by students with significant disabilities. *Education and Training in Developmental Disabilities, 42*, 339-352.
- Bostow, D., & Geiger, O. (1976). Good behavior game: A replication and systematic analysis with a second grade class. *School Applications of Learning Theory, 8*, 18-27.
- Browder, D., Liberty, K., Heller, M., & D'Huyvetters, K. (1986). Self-management by teachers: Improving instructional decision making. *Professional School Psychology, 1*, 165-175.
- Burg, M., Reid, D., & Lattimore, J. (1979). Use of a self-recording and supervision program to change institutional staff behavior. *Journal of Applied Behavior Analysis, 12*, 363-375.
- Cates, G. (2005). A review of the effects of interspersing procedures on the stages of academic skill development. *Journal of Behavioral Education, 14*, 305-325. doi: 10.1007/s10864-005-8652-8

- Codding, R. S., Feinberg, A. B., Dunn, E., & Pace, G. M. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis, 38*, 205-219.
- Codding, R., Livanis, A., Pace, G., & Vaca, L. (2008). Using performance feedback to improve treatment integrity of classwide behavior plans: An investigation of observer reactivity. *Journal of Applied Behavior Analysis, 41*, 417-422.
- Cooper, J., Heron, T., & Heward, W. (2007). *Applied behavior analysis* (2nd ed). Upper Saddle River, NJ: Pearson.
- Darveaux, D. (1984). The Good behavior game plus merit: Controlling disruptive behavior and improving student motivation. *School Psychology Review, 13*, 510-515.
- DiGennaro, F., Martens, B., & McIntyre, L. (2005). Increasing treatment integrity through negative reinforcement: Effects on teacher and student behavior. *School Psychology Review, 34*, 220-231.
- Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., Laudolff, J., Turkan, J. S., Ford, C., & Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology, 14*, 317-345.
- Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350. doi: 10.1007/s10464-008-9165-0

- 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.
- Embry, D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*, 373-297.
- Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, R. (2008). *Reducing behavior problems in the elementary school classroom: A practice guide* (NCEE#2008-012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institutes of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/wwc/publications/practiceguides>
- Evidence-Based Intervention Work Group, The (2005). Theories of change and adoption of innovations: The evolving evidence-based intervention and practice movement in school psychology. *Psychology in the School, 42*, 475-494. doi: 10.1002/pits.20086
- Fishbein, J. & Wasik, B. (1981). Effects of the good behavior game on disruptive library behavior. *Journal of Applied Behavior Analysis, 14*, 89-93.
- Florida Department of Education. (2007). *Florida performance pay guidance: 2007–2008 and beyond*. Tallahassee, FL: Author. Retrieved from <http://www.fldoe.org/PerformancePay/pdfs/MeritAwardProgram.pdf>
- Gilbertson, D., Witt, J., Singletary, 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. doi: 10.1007/s10864-007-9043-0

- Gottfredson, D., Gottfredson, G., & Hybl, L. (1993). Managing adolescent behavior: A multi-year, multi-school study. *American Educational Research Journal, 30*, 179-215.
- Goe, L., Bell, C., & Little, O. (2008). Approaches to evaluating teacher effectiveness: A research synthesis. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from <http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf>
- Gravina, N., Austin, J., Schoedtder, & Loewy, S. (2008). The effects of self-monitoring on safe posture performance. *Journal of Organizational Behavior Management, 28*, 238-259. doi: 10.1080/01608060802454825
- Gresham, F. (1989). The assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review, 18*, 37-50.
- Gresham, F., Gansle, K., & Noell, G. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis, 26*, 257-263.
- Hagermoser Sanetti, L., Gritter, K., & Dobey, L. (2011). Treatment integrity of interventions with children in the school psychology literature from 1995-2008. *School Psychology Review, 40*, 72-84.
- Hagermoser Sanetti, L., & Kratochwill, T. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review, 38*, 445-459.
- Hagermoser Sanetti, L., Luiselli, J., & Handler, M. (2007). Effects of verbal and graphic performance feedback on behavior support plan implementation in a public elementary school. *Behavior Modification, 31*, 454-465.

- Hairrell, A., Rupley, W., Edmonds, M., Larsen, R., Simmons, D., Wilson, V., et al. (2011). Examining the impact of teacher quality on fourth-grade students' comprehension and content-area achievement. *Reading & Writing Quarterly*, 27, 239-260. doi: 10.1080/10573569.2011.560486
- Han, S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology*, 33, 665-679. doi: 10.1007/s10802-0025-7646-2
- Han, S., & Weiss, B. (2008). Maintaining program fidelity after the thrill—and external support—is gone. In S. Evans, M. Weist, & Z. Serpell (Eds.), *Advances in school-based mental health interventions: Best practices and program models, Vol II*. (pp. 1-17). Kingston, NJ: Civic Research Institute.
- Harris, V., & Sherman, J. (1973). Use and analysis of the “Good Behavior Game” to reduce disruptive classroom behavior. *Journal of Applied Behavior Analysis*, 6, 405-417.
- Hoover, N.L., & Carroll, R.G. (1987). Self-assessment of classroom instruction: An effective approach to inservice education. *Teaching and Teacher Education*, 3, 179-191.
- Idol, L. (1993). *Special educator's consultation handbook* (2nd ed.). Austin, TX: PRO-ED.
- Johnson, M. R., Turner, P. F., & Konarski, E. A. (1978). The Good Behavior Game: A systematic replication in two unruly transitional classrooms. *Education and Treatment of Children*, 1, 25-33.
- Jones, K., Wickstrom, K., & Friman, P. (1997). The effects of observational feedback on treatment integrity in school-based behavioral consultation. *School Psychology Quarterly*, 12, 316-326.

- Kalis, T., Vannest, K., & Parker, R. (2007). Praise counts: Using self-monitoring to increase effective teaching practices. *Preventing School Failure, 51*, 20-27.
- Kellam, S., Ling, X., Merisca, R., Hendricks Brown, C., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology, 10*, 165-185.
- Keller, C., Brady, M., & Taylor, R. (2005). Using self evaluation to improve student teacher interns' use of specific praise. *Education and Training in Developmental Disabilities, 40*, 368-376.
- Kennedy, C. (2005). *Single-case designs for educational research*. Boston: Allyn and Bacon.
- Kilbourn, B. (1991). Self-monitoring in teaching. *American Educational Research Journal, 28*, 721-736. Retrieved from <http://www.jstor.org/stable/1163018>
- Kissel, R., Whitman, T., & Reid, D. (1983). An institutional staff training and self-management program for developing multiple self-care skills in severely/profoundly retarded individuals. *Journal of Applied Behavior Analysis, 16*, 395-415.
- Kleinman, K. & Saigh, P. (2011). The effects of the Good Behavior Game on the conduct of regular education New York City high school students. *Behavior Modification, 35*, 95-105.
- Kuhn, M., & Stahl, S. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*, 3-21. doi: 0.1037/0022-0663.95.1.3

- Lane, K. L., Bocian, K. A., 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.
- Lannie, A., & McCurdy, B. (2007). Preventing disruptive behavior in the urban classroom: The effects of the Good Behavior Game on student and teacher behavior. *Education and Treatment of Children, 30*, 85-98.
- Martens, B., Hiralall, A., & Bradley, T. (1997). A note to teacher: Improving student behavior through goal setting and feedback. *School Psychology Quarterly, 12*, 33-41.
- Mathers, C., Olivia, M., & Lane, S. (2008). Improving instruction through effective teacher evaluation: Options for states and districts. Washington, DC: National Comprehensive Center for Teacher Quality.
- McCurdy, B., Lannie, A., & Barnabas, E. (2009). Reducing disruptive behavior in an urban school cafeteria: An extension of the Good Behavior Game. *Journal of School Psychology, 47*, 39-54. doi:10.1016/j.jsp.2008.09.003
- McGoey, K., Schneider, D., Rezzetano, K., Prodan, T., & Tankersley, M. (2010). Classwide intervention to manage disruptive behavior in the kindergarten classroom, *Journal of Applied School Psychology, 26*, 247-261. doi: 10.1080/15377903.2010.495916
- McIntyre, L., Gresham, F., DiGennaro, F., & Reed, D. (2007). Treatment integrity of school-based interventions with children in the Journal of Applied Behavior Analysis 1991-2005. *Journal of Applied Behavior Analysis, 40*, 659-672.

- Medland, M., & Stachnik, T. (1972). Good-behavior game: A replication and systematic analysis. *Journal of Applied Behavior Analysis*, 5, 45-51.
- Mercer, C. D., Mercer, A. R., Pullen, P. C. (2011). *Teaching students with learning problems* (8th ed.). Upper Saddle River, NJ: Pearson.
- Mooney, P., Epstein, M., Reid, R., & Nelson, J. (2003). Status and trends of academic intervention research for students with emotional disturbance. *Remedial and Special Education*, 24, 273–287.
- Morgan, M. (1987). Self-monitoring and goal setting in private study. *Contemporary Educational Psychology*, 12, 1-6.
- Mortenson, B., & Witt, J. (1998). The use of weekly performance feedback to increase teacher implementation of a preferral academic intervention. *School Psychology Review*, 27, 613-627.
- Noell, G., Duhon, G., Gatti, S., & Connell, J. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review*, 31, 217-234.
- Noell, G., Witt, J., Gilbertson, D., Ranier, D., & Freeland, J. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly*, 12, 77-88.
- Noell, G., Witt, J., LaFleur, L., Mortenson, B., Ranier, 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., Witt, J., Slider, N., Connell, J. Gatti, L., Williams K., et al. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review*, 34, 87-106.

- Peterson, L., Homer, A., & Wonderlich S. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis, 15*, 477-492.
- Petscher, E. S., & Bailey, J. S. (2006). Effects of training, prompting, and self-monitoring on staff behavior in a classroom for students with disabilities. *Journal of Applied Behavior Analysis, 39*, 215-226.
- Reinke, W., Lewis-Palmer, T., & Merrell, K. (2008). The classroom check-up: A classwide teacher consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review, 37*, 315-332.
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes* (Report). Washington, DC: Economic Policy Institute.
- Ruiz-Olivares, R., Pino, M., & Herruzo, J. (2010). Reduction of disruptive behaviors using an intervention based on the Good Behavior Game and the Say-D0-Report correspondence. *Psychology in the Schools, 47*, 1046-1058. doi: 10.1002/pits.20523
- Saigh, P., & Umar, A. (1983). The effects of a Good Behavior Game on the disruptive behavior of Sudanese elementary school students. *Journal of Applied Behavior Analysis, 16*, 339-344.
- Salend, S. J., Reynolds, C. J., & Coyle, E. M. (1989). Individualizing the Good Behavior Game across types and frequency of behavior with emotionally disturbed adolescents. *Behavior Modification, 13*, 108-126.
- Sanetti, L. & Kratochwill, T. (2008). Treatment integrity in behavioral consultation: Measurement, promotion, and outcomes. *International Journal of Behavioral Consultation and Therapy, 4*, 95-114.
- Scheeler, M., Ruhl, K., & McAfee, J. (2004). Providing performance feedback to teachers: A review. *Teacher Education and Special Education, 27*, 396-407.

- Shake, M. (1986). Teacher interruptions during oral reading instruction: Self-monitoring as an impetus for change in corrective feedback. *Remedial and Special Education, 7*, 18-24.
- Smith, D. D. (1981). *Teaching the learning disabled*. Upper Saddle River, NJ: Prentice Hall.
- Smith, S. W., Daunic, A. P., & Taylor, G. G. (2007). Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. *Education and Treatment of Children, 30*, 121-134.
- Snell, M. et al. (2010). Twenty years of communication intervention research with individuals who have severe intellectual and developmental disabilities. *American Association on Intellectual and Developmental Disabilities, 115*, 364-380. doi: 10.1352/1944-7558-115-5364
- Stichter, J., Lewis, T., Richter, M., Johnson, N., & Bradley, L. (2006). Assessing antecedent variables: The effects of instructional variables on student outcomes through in-service and peer coaching professional development models. *Education and Treatment of Children, 29*, 665-692.
- Sulzer-Azaroff, B., & Mayer, G. (1991). *Behavior analysis for lasting change*. Fort Worth, TX: Harcourt Brace College.
- Sutherland, K., & Wehby, J. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *The Journal of Special Education, 35*, 161-171.

- Swiezy, N., Matson, J., & Box, P. (1993). The Good Behavior Game: A token reinforcement system for preschoolers. *Child & Family Behavior Therapy, 14*, 21-32.
- Tankersley, M. (1995). A group-oriented contingency management program: A review of research on the Good Behavior Game and implications for teachers. *Preventing School Failure, 40*, 19-24.
- Telzrow, C., McNamara, K., & Hollinger, C. (2000). Fidelity of problem-solving implementation and relationships to student performance. *School Psychology Review, 29*, 443-461.
- Warner, S. P., Miller, F. D., & Cohen, M. W. (1977). Relative effectiveness of teacher attention and the “good behavior game” in modifying disruptive classroom behavior. *Journal of Applied Behavior Analysis, 10*, 737.
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.
- Witt, J., Noell, G., LaFleur, & Mortenson, B. (1997). Teacher use of interventions in general education settings: Measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis, 30*, 693-696.
- Yeaton, W., & 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.