

ENHANCING TREATMENT INTEGRITY THROUGH  
SYSTEMATIC FADING WITH INDISCRIMINABLE  
CONTINGENCIES

By

THOMAS J. GROSS

Bachelor of Arts in Psychology  
Marquette University  
Milwaukee, WI  
2002

Master of Arts in Community Counseling  
Marquette University  
Milwaukee, WI  
2004

Submitted to the Faculty of the  
Graduate College of the  
Oklahoma State University  
in partial fulfillment of  
the requirements for  
the Degree of  
DOCTOR OF PHILOSOPHY  
August, 2013

ENHANCING TREATMENT INTEGRITY THROUGH  
SYSTEMATIC FADING WITH INDISCRIMINABLE  
CONTINGENCIES

Dissertation Approved:

Gary Duhon, Ph.D.

---

Dissertation Adviser

Terry Stinnett, Ph.D.

---

Brian Poncy, Ph.D.

---

C. Robert Davis, Ph.D.

---

Name: THOMAS J. GROSS

Date of Degree: DECEMBER, 2013

Title of Study: ENHANCING TREATMENT INTEGRITY THROUGH SYSTEMATIC  
FADING WITH INDISCRIMINABLE CONTINGENCIES

Major Field: EDUCATIONAL PSYCHOLOGY, SPECIALIZATION IN SCHOOL  
PSYCHOLOGY

ABSTRACT: Frequently school psychologists are asked to develop treatments for teachers to remediate students' academic skills or social behaviors. When teachers implement them with high levels of treatment integrity they are beneficial to students. Treatment integrity has been increased by using direct training, performance feedback, and negative reinforcement procedures. Still, treatment integrity maintenance has not received the same amount of systematic inquiry. This study investigated treatment integrity maintenance over time after a systemic fading procedure with indiscriminable contingencies was implemented. The results showed that the teachers continued to implement treatment protocols at a high level after performance feedback was faded out. Students also fared better when treatment integrity was high.

## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
II. REVIEW OF LITERATURE.....	9
Treatment Integrity.....	9
Treatment Integrity.....	9
Measuring Treatment Integrity Through Adherence.....	10
Teacher Treatment Implementation with Consultation.....	12
Indirect and Direct Assessment of Treatment Integrity.....	13
Accountability.....	18
Role of the Consultant.....	18
Teacher Training.....	19
Performance Feedback Emphasis.....	24
Maintenance of Treatment integrity.....	26
Issues in Performance Feedback Literature.....	26
Revelations in Maintenance Phases.....	27
Fading Procedures.....	31
Indiscriminable Contingencies.....	36
Indiscriminable Contingencies.....	36
Examples in Student Interventions.....	37
Application to Teacher Treatment Integrity.....	39
Prereferral Intervention Teams.....	40
Multiple Baseline Designs.....	42
III. METHODOLOGY.....	46
Overview of the Student Assistance Team.....	46
Overview of Procedures.....	46
Participants.....	47
Setting.....	48
Materials.....	49
Problem Assessment Measures.....	49
Intervention Materials.....	50
Dependent Measures.....	52
Treatment Integrity.....	52
Student Outcomes.....	53
Experimental Design.....	53

	Direct Training Procedures.....	54
	Implementation Baseline.....	54
	Performance Feedback with Direct Rehearsal .....	55
	Fading 1 .....	55
	Fading 2.....	56
Chapter		Page
	Fading 3.....	56
	Maintenance.....	57
IV.	FINDINGS .....	31
	Treatment Integrity .....	58
	Comparison of Treatment Integrity and Student Performance .....	59
	Student Outcomes .....	60
	Comparison of Teacher and Student Data.....	61
V.	CONCLUSION .....	63
	Treatment Integrity .....	63
	Use of Indiscriminable Contingencies.....	64
	Impact on Student Outcomes .....	65
	Contributions.....	65
	Limitations .....	67
REFERENCES.....		69
APPENDICES.....		77

## LIST OF TABLES

Table	Page
1	89
2	90

## LIST OF FIGURES

Figure	Page
1	78
2	79
3	80
4	81
5	82
6	83
7	84
8	85
9	86
10	87
11	88

## CHAPTER I

### INTRODUCTION

Federal mandates, including the No Child Left Behind Act of 2001 and the more recent American Recovery and Reinvestment Act of 2009, emphasize continual improvement in children's academic competence. Generally, students respond better to early remediation (Shaywitz, Morris & Shaywitz, 2008), making it critical to develop and implement valid early interventions (Burke & Hagan-Burke, 2007). School psychologists are asked to develop interventions with teachers and teachers are frequently asked to implement them (Sterling-Turner, Watson & Moore, 2002). Treatment integrity is the degree that an intervention plan is implemented with accuracy and consistency (Gresham, 1989; Gresham, 2005). This is often measured as adherence, which is the number of components from a specific treatment protocol that are implemented and recorded through self-report, observation (Griffith, Duppong-Hurley & Hagan, 2009) or permanent products (Noell & Witt, 1999) following consultation. Additionally, treatment integrity is believed to be necessary for demonstrating a functional relationship between intervention procedures and changes in behaviors (Gresham, 2005). Assuring specific procedures are followed, is a key piece in determining if the intervention created an observed change.



Treatment integrity is detected through indirect measures, manualized treatments or direct measures. Self-monitoring is used as a cost effective, indirect rating of integrity. However, when teachers complete self-reports they tend to over-estimate their integrity (Smith, Daunic, & Taylor, 2007; Wickstrom, Jones, LaFleur, & Witt, 1998). Nonetheless, when they are monitored by an observer they give more accurate self-assessments (Hagermoser-Sanetti & Kratochwill, 2009). One way to resolve this is through the use of manualized treatments. Manualized treatments are written guidelines for the implementation of acceptable interventions, which give the characteristics and sequence of the treatment process, standardize treatment implementation, and train the implementers (Moncher & Prinz, 1991). They provide a system of intervention that allows for self-assessment and provides the guidance and clarification of a supportive consultant (Smith et al., 2007). Still, there is no independent verification of the treatment being used as directed (Gresham, 2005). Permanent products are references that are generated from the completion of each treatment component. These appear to be more accurate than other indirect measures of integrity (Gresham, MacMillian, Beebe-Frankeberger, & Bocian, 2000). Direct measures of integrity include systemic observations in a naturally occurring or analogous context, videotaping, and audio recording. For either direct measures or permanent products to be effective, key elements of treatment must be clearly identified and subtasks must be specifically defined behaviors, measures, timing, and distances. Secondly, the demonstration or omission of behaviors should be assessed to provide the integrity over an observation session. Lastly, the percentage of components implemented by consultees can be calculated over time (Gresham, 1989; Gresham et al., 2000).

Assuring treatment integrity is commonly agreed to improve outcomes for students. Researchers contend that higher rates of integrity correspond with raised levels of positive behavior change in children, while low levels of integrity correspond with decreases in expected child behaviors (Reinke, Lewis-Palmer, & Merrell, 2008). Further it is found that high levels of integrity are related to decreases in inappropriate behaviors, remaining effects after treatment, and better outcomes, independent of the consultation type (Stage et al., 2008). Subsequent reviews of the literature have found moderate relationships between treatment integrity and treatment outcome (Gresham, 2005). More recently, the use of applied behavior analysis has demonstrated the correspondence between treatment integrity and intervention effectiveness (Wood, Umbreit, Liaupsin, & Gresham, 2007). While it is the teachers' responsibility to implement treatment with integrity, it can be viewed as the consultant's responsibility to ensure integrity is monitored and maintained. Providing recommendations alone is insufficient to guarantee adherence to an intervention protocol and is unsatisfying to way to draw conclusions about the effectiveness of interventions (Noell, 2008; Noell & Witt, 1996). We must encourage sustained teacher efforts through means such as direct training and performance feedback for intervention implementation (Noell, 2008).

Direct training of teachers to implement specific teaching behaviors has proved fruitful for providing correct and consistent performance of teacher tasks. Direct rehearsal with error correction of specific behaviors has been found to increase a broad range of specific teacher behaviors in the classroom (Ward, Johnson, & Konukman, 1998). Further, a comparison between didactic training (DT), modeling training (MT), and rehearsal feedback training (RFT) found that RFT was superior to both DT and MT (Sterling-Turner, Watson, Wildmon, Watkins, & Little, 2001). The implication is that indirect training procedures are

insufficient to adequately increase teacher behaviors associated with treatment integrity. Consultants must train teachers through instruction, rehearsal and performance feedback to ensure proper implementation of treatment procedures. Many other studies lend support to this contention (Coddington, Feinberg, Dunn, & Pace, 2005; Gilbertson, Witt, LeFleur-Singletary, & VanDerHeyden, 2007; Noell, Witt, Gilbertson, Ranier, & Freedland, 1997; Witt, Noell, LeFleur, & Mortenson, 1997). It has been proposed that DT procedures will result in less integrity and therefore less change in student behaviors, while RFT will result in more integrity and therefore more change in student behaviors. Research using teacher student dyads has supported this contention. RFT demonstrated increased implementation of treatment components along with suppression of undesirable student behaviors (e.g. out of seat, inappropriate vocalizations) and increases in desired behaviors (e.g. compliance). Additionally, it was emphasized that beginning with DT procedures may promote the absence of behavior change and negative perceptions of the consultation process (Sterling-Turner et al., 2002).

Maintenance of behaviors taught to teachers has long been of concern for consultants. The impact of modeling, practice and feedback has shown some promise for demonstration of acquired teaching behaviors and fluency in delivery of reading programming after termination of training at five and twenty sessions (Carnine & Fink, 1978). Others have found that behaviors are maintained at five, ten, and fifteen week follow-up integrity checks (Coddington et al., 2005). This was following eight to 22 weeks of performance feedback; however, the authors cautioned that maintenance scores may be inflated due to reactivity of being observed on the part of the teacher, and continued performance feedback may be necessary for persistent treatment integrity (Coddington et al., 2005). Some research has shown

the maintenance of integrity to be highly variable across teachers, as well as, for individual teachers (Gilbertson et al., 2007; Noell et al., 1997; Noell et al., 2000; Witt et al., 1997).

What is surprising about many studies that include maintenance data is that they are often only a few points, usually two or three, and fail to demonstrate clear trends in teacher adherence to treatment protocols. However, there is some indication that after performance feedback is removed the behaviors remain, but follow a downward trend (Reinke et al., 2008), or once public feedback through team meetings is removed teachers tend to fail to adhere to intervention protocols (Duhon, Mesmer, Gregerson, & Witt, 2009). It appears that teachers can retain behaviors specific to interventions, but the withdrawal of support for teachers to maintain high adherence to the intervention works to the detriment of treatment integrity.

To combat drop-offs in treatment integrity and conserve time resources, researchers have added fading schedules of performance feedback into their designs. One study reduced feedback session from daily to every other day when teachers achieved 100% integrity for four consecutive days. Of the five teacher-participants, three received the reduction in feedback sessions. Data collected for one teacher only covered three days after, but showed continued effect. However, the other two teachers demonstrated some variability over the course of five school days (Noell et al., 2000). Further research findings indicate that performance feedback remains somewhat effective if it is faded from daily to once every three days. Nonetheless, teachers tend to have greater variability in treatment integrity and may need a more structured performance feedback component (Noell, Duhon, Gatti, & Connell, 2002).

Digennaro, Martens, and Kleinmann (2007) sought to examine teachers' treatment integrity after goal setting, after goal setting with performance feedback, determine if treatment integrity remained high after thinning performance feedback from daily to every other week, and demonstrate a relationship between treatment integrity and student behavior. The researchers measured integrity across six phases: (1) pre-training base-line, (2) training, (3) implementation baseline, (4) goal setting and student performance feedback, (5) teacher performance feedback and direct rehearsal with meeting cancellation, and (6) fading. The fading procedure kept in place performance feedback and direct rehearsal procedures, but aimed to reduce them from daily to every other day, once per week and then once every two weeks. The thinning of performance feedback was planned to take place after the teacher completed three consecutive sessions with 100% integrity. Consistent with prior research it was found that performance feedback with direct rehearsal and meeting cancellation had the greatest effect for increasing integrity, reducing variability in adherence and decreasing undesirable behaviors in students. It is proposed that meeting avoidance has acts to negatively reinforce the teachers' treatment adherence behaviors (Digennaro, Martens & McIntyre, 2005; Digennaro et al., 2007). Nevertheless, only one of the four teachers in the study received the thinning to desired levels. This teacher demonstrated only slight variability in integrity and the level of student off-task behavior remained lower than at baseline. For the remaining three teachers, the schedules were not thinned due to constraints of the school calendar. There data was at various stages of fading and showed differing degrees of variability in treatment integrity. This has spurred the authors to identify fading procedures as an important area for further investigation (Digennaro et al., 2007).

Prior research with students has identified maintenance as the durability of behavior as treatment has been removed (Freeland & Noell, 2002). Importantly, treatment integrity research seeks to provide treatment to the teacher through performance feedback. The fading procedures used attempts to establish consistent demonstration of behavior sets as feedback is removed. What may add to the fading procedures is the use of indiscriminable contingencies to help increase intervention protocol adherence. Central to this is that concurrent use of intermittent and delayed reinforcement may make it difficult for people to discriminate a shift from a treatment phase to thinned schedules, and contributes to substantial maintenance of responding even after reinforcement is discontinued (Freeland & Noell, 2002). As in the previous study by Digennaro et al. (2007) the use of fading from daily rehearsal and feedback meetings to every other week can be viewed as increasing time between reinforcement (changing intermittent negative reinforcement), while delaying information about reinforcement to the teacher until well after the intervention is completed (increasing delay before reinforcement). Had the researchers been able to complete the fading procedure, they could have demonstrated an effective maintenance procedure. Nonetheless, it may also be beneficial for treatment integrity to make indiscriminable contingencies salient to the intervening teacher. They will know what procedures and reinforcers to expect, but they would still have difficulty discriminating the treatment phase from the maintenance phase.

Given the preponderance of evidence for the efficacy of direct training procedures with performance feedback it seems unnecessary to scrutinize its efficacy. Rather, it is more beneficial to examine procedures to maintain treatment integrity. Specifically, the use of fading procedures in conjunction with indiscriminable contingencies should be fruitful to this

end. One goal of this study is to demonstrate that treatment integrity can be maintained as the scheduled performance feedback is reduced. A second goal is to demonstrate that indiscriminable contingencies can be effectively integrated into interventions for increasing teacher adherence to treatment protocols. Lastly, it should be demonstrated that treatment integrity and its maintenance are associated with more desirable student outcomes.

## CHAPTER II

### REVIEW OF THE LITERATURE

#### **Treatment Integrity**

**Treatment Integrity.** Treatment integrity is broadly conceptualized as implementation of an intervention as intended (Gresham, 2005; Perepletchikova & Kazdin, 2005), or as designed by a consultant (Gresham, 1989). This includes assessing the accuracy and consistency through monitoring the delivery of each component (Smith, Daunic & Taylor, 2007). Assuring high levels of treatment integrity is associated with changes in behavior due to treatment effects, outcomes for students or clients are a result of planned changes to the environment, providing evidence that an intervention worked, and demonstrations of the internal and external validity of an intervention (Gresham, 1989; Perepletchikova & Kazdin, 2005; Smith et al., 2007; Griffith, Hurley & Hagaman, 2009). It is a necessary element to demonstrate a functional relationship between a therapeutic procedure and a change in behavior; however, it is not sufficient in and of itself (Gresham, 2005).

Treatment integrity is composed of three components: (1) adherence, (2) competence, and (3) treatment differentiation. Adherence is the degree to which a procedure is carried out as planned (Perepletchikova & Kazdin, 2005), as defined by



the operational definitions of the protocol (Griffith et al., 2009). Poor adherence can be understood as the application of an all together different treatment and leaves interpretation of the intended treatment's effect ambiguous (Perepletchikova & Kazdin, 2005). Additionally, poor adherence is associated with a lack of evidence-based practices in education (Smith et al., 2007), a failure to rule out extemporaneous variables as the cause for behavior change, and ineffective consultation practices (Gresham, 1989).

Key to adherence is monitoring treatment delivery through clearly specified methods to establish treatment integrity (Smith et al., 2007). The accuracy of treatment delivery is most often determined by calculating the within session completion of steps. Operationally stated, the number of treatment steps completed is divided by the total number of treatment steps and multiplied by 100% to provide the percentage of treatment steps followed (Gresham, 1989; Smith et al., 2007). While this method provides us with an empirical basis for the degree of adherence demonstrated, there is no empirically established acceptable level. Some have argued that there needs to be a definitive standard to assure accuracy in implementation and corresponding measures (Gresham, 2005). Nonetheless, it is widely acknowledged that 80% or greater adherence is a high level of treatment integrity (Smith et al., 2007).

**Measuring Treatment Integrity through Adherence.** Methods to assess adherence to treatment protocol include: direct observation, feedback, self-monitoring, treatment manualization, and permanent products. Direct observation is when an observer indicates on a protocol if the interventionist presented a treatment component or not. While this allows for a high degree of confidence in adherence, it is costly, which may prohibit its use in many situations (Smith et al., 2007). Feedback is when observation is

combined with feedback during implementation. This method has similarly cost prohibitive if done at the time of the session (Smith et al., 2007). Self-monitoring is when the treatment provider reports the steps he has completed or not (Smith et al., 2007). This is an enticing option because it is cost and time efficient; however, there is evidence that self-reported adherence tends to significantly inflate the actual steps completed (Smith et al., 2007; Wickstrom, Jones, LaFleur, & Witt, 1998). Manualized treatments are when a trainer provides a reference to the treatment provider for the delivery of an intervention (Smith et al., 2007). Manualized treatments should increase treatment integrity, provided there is adequate training (Perepletchikova & Kazdin, 2005), but one shortcoming is that there is no way to independently verify if the treatment is implemented as intended (Gresham, 2005). Lastly, the evaluation of permanent products looks for byproducts of an intervention to evaluate the content of the treatment session (Smith et al., 2007). This method has provided accurate rates of treatment integrity throughout an intervention and works well in synthesis with performance feedback (Gresham, MacMillian, Beebe-Frankeberger & Bocian, 2000; Noell & Witt, 1999). Still, it also highlights that a combination of methods should be used to measure adherence and assure treatment integrity (Smith et al., 2007).

Besides analyzing the validity of the intervention being used, treatment adherence data is important to inform consultants of when to intervene upon the treatment provider, to increase treatment integrity. Some have dedicated a great deal of energy to determine if acceptability of an intervention impacts integrity. Repeatedly, the correlation between high acceptability and high integrity has failed to be significant (Gresham, 2005; Noell &

Witt, 1999). Additionally, making the intervention process public has not guaranteed integrity, as contingencies for treatment integrity are rarely employed (Gresham, 1989). Other avenues have developed promising methods to enhance treatment integrity, such as indirect training, direct training and performance feedback. Indirect training consists of didactic instructions, written materials regarding the interventions rationale, scripts, and check list for tasks. This method may take a good deal of extensive training to assure integrity (Perepletchikova & Kazdin, 2005). Direct training involves practicing the intervention through role-plays, modeling, feedback, rehearsal and follow-up sessions. This is believed to promote accurate implementation of a treatment (Perepletchikova & Kazdin, 2005). Performance feedback has resulted in far better results for treatment implementation and the resulting behavioral outcomes than either indirect or direct training alone (Gresham, 2005). Moreover, performance feedback fits well with long held recommendations for increasing integrity, such as plotting integrity data, reviewing the plan with the consultee, and monitoring a plan to give corrective feedback (Gresham, 1989). Performance feedback has also shown efficacy in increasing treatment integrity after it has declined (Noell & Witt, 1999).

**Teacher Treatment Implementation with Consultation.** While there is an abundance of information about methods to improve treatment integrity, little research has examined the extent to which teachers implement interventions developed within the consultation process (Noell & Witt, 1999). Consultation is a major device of psychologists in schools and it seems disheartening that few resources and procedures to ensure and maintain treatment integrity are identified. Additionally, research has not looked at how examination of the educational context and treatment integrity interact

(Noell & Witt, 1999). Following another review of the literature regarding integrity in the consultation process, Noell and Witt (1999) reexamined the articles and found that of the original 46 studies, 11 assessed consultation practices. Two studies provided no information about implementation, three studies used direct observation or permanent products, and six used indirect measures (e.g. self-report, ratings by research team members; Noell & Witt, 1999).

A recent review of the literature in emotionally or behaviorally disturbed students' treatment looked at adherence to intervention protocols in 44 articles, spanning from 1965 to 2005 (Griffith et al., 2009). Of these articles, 23 reported integrity data, two mentioned integrity data was collected, and 19 reported no integrity data (Griffith et al., 2009). Low levels of integrity examination are found in learning disabilities intervention research as well (Gresham et al., 2000). It is then reasonable to assume that there has been little empirical investigation of the procedures used to track and maintain teacher treatment integrity. Further, the reported treatment integrity data is variable in how it is measured across studies and yields variable results across studies (Griffith et al., 2009; Moncher & Prinz, 1991; Noell & Witt, 1999). This appears to call for providing ways to systematize recording treatment adherence and increase the consistency of adherence to intervention protocols for teachers.

### **Indirect and Direct Assessment of Teacher Integrity**

The need to systematically record treatment integrity data and elevate levels of teacher treatment integrity is echoed throughout the literature. Teachers' treatment integrity for student behavior with behavior problems was examined in a comparison of interventions designed through functional behavioral assessment (FBA), and behavioral

consultation (BC; Stage et al., 2008). Behavioral observations were used to determine if teachers delivered all of the components of the behavior plan established through either FBA or BC. Regardless of the method used for designing the intervention, FBA or BC, treatment integrity had the greatest impact on child outcomes. Follow-up interviews with the teachers demonstrated that they found the interventions useful, but would have liked more support for implementing the procedures (Stage et al., 2008). This indicates that support for high treatment integrity has the potential to increase favorable outcomes for students.

There are many ways researchers have attempted to increase treatment integrity, measure treatment adherence, and enhance integrity for interventions used by teachers. Yet, our understanding of this is limited in naturalistic contexts and is far less understood than intervention components. Wickstrom, Jones, LaFleur, and Witt (1998) acknowledged that the understanding of a treatment's impacts on behavior out-paced the understanding of the understanding of their application. They proposed to assess the use of treatment protocols within the naturalistic environment of elementary schools with teachers as interventionists. The researchers acted as consultants for three week. To measure treatment integrity, they employed three measures: (1) the baseline and intervention rating form (BIRF) required teachers to monitor child behavior and their responses; (2) independent observations of the presence of a stimulus product associated with the intervention were taken; (3) direct observations using a partial interval recording system were used to record if the teacher used the intervention stimulus product (Wickstrom et al., 1998).

The mean BIRF integrity across teachers was 54%, with a mean of 67% at baseline and 41% during the final week of consultation (Wickstrom et al., 1998). The teachers' self report of integrity indicated they completed the respective interventions about half of the time. During the beginning of consultation they reported adhering to the protocol more, but decreased their delivery of treatment as prescribed over the course of three weeks. The observations of the stimulus products present measured treatment adherence at 62% (Wickstrom et al., 1998). This indicated that the completed materials for the interventions were present at low levels. Lastly, the direct observations of teachers yielded a mean of 4% for treatment integrity. This meant the teachers were delivering the intervention for one out of every 20 target behaviors (Wickstrom et al., 1998).

The researchers uncovered that the reliance on indirect measures may be risky in the assessment of treatment integrity (Wickstrom et al., 1998). When examining the results of all three integrity measures, it appears that direct observation of teacher implementation provided the most accurate picture, whereas less rigorous measures maybe misleading. The observed average teacher integrity implies a great deal of uncertainty for the delivery of intervention steps. This is further complicated by self-reports indicating that the teachers may not sustain interventions or the course of a few weeks.

A grim outlook for the use of self-report has not deterred proponents of the practice from developing systems to put it into practice. Self-report (or self-monitoring) is an enticing option because it is cost and time effective (Smith et al., 2007). The Treatment Integrity Planning Protocol (TIPP) is a standardized three-stage process for developing plans and self-assessments of treatment integrity for school-based

interventions (Hagermoser-Sanetti & Kratochwill, 2009). In evaluating the TIPP as a valid process for increasing treatment integrity, it was hypothesized that combing planning and self-evaluation would increase and maintain integrity. Treatment integrity was measured by calculating the number of permanent products and use of the TIPP, either with or without consultant assistance (Hagermoser-Sanetti & Kratochwill, 2009).

The implementation of the TIPP was associated with increased treatment integrity and decreased variability in treatment adherence, as measured by permanent products (Hagermoser-Sanetti & Kratochwill, 2009). Nonetheless, while the TIPP purports to be a self-report measure, it more closely resembles being an intervention to assure treatment integrity. It demonstrated the greatest effect when tied to meetings with a consultant, a consultant aided in the design of the treatment check sheet, and a consultant directed the teacher to complete the form every day. The option to just use the TIPP manual, which is analogous to solely relying on self-report, was fruitless. Further, maintenance of adherence was not measured (Hagermoser-Sanetti & Kratochwill, 2009). It appears that accuracy in self-report needs to be trained in and coaxed from teachers, and there is no data to tell us if this effort has a lasting effect on treatment integrity.

Direct measures of integrity allow for the consultant to determine if poor results are due to the intervention or adherence in implementation. In an attempt to demonstrate the correspondence of treatment effect and treatment integrity, student on-task behavior was charted and then observations of treatment adherence were overlaid on student performance (Wood, Umbreit, Liaupsin & Gresham, 2007). Remarkable, the data points of student on-task behavior nearly resembled the data points for teacher treatment integrity. When the teacher implementation was correct, student on-task behaviors

reached 90% for timed intervals, while incorrect implementation resulted in 10% on-task behaviors for timed intervals (Wood et al., 2007).

Research efforts should not discard self-report all together; rather its worth can be seen in light of other methods to enhance treatment integrity. The Classroom Check-Up (CCU) is a system of consultation, meant to enhance treatment implementation through motivating the teacher. It consists of five steps, which employ assessment, feedback regarding the assessment, collaboration to develop options, intervention planning, and self-monitoring for treatment integrity (Reinke, Lewis-Palmer & Merrell, 2008). One study hypothesized that the use of the CCU with self-monitoring and visual feedback would be superior to the CCU with self-monitoring only. Teachers' rates of praise for specified student behaviors were targeted. Visual feedback was given in the form of a graph depicting the teachers' rates of praise, which was linked to direct measures of teacher behaviors within the classroom (Reinke et al., 2008).

Teacher praise for specific behaviors and student disruptions were monitored over baseline and CCU with self-monitoring, CCU with self-monitoring and visual performance feedback, and maintenance phases. Baseline data was collected before consultation began and maintenance data was collected one month after visual performance feedback was withdrawn (Reinke et al., 2008). The results indicated that self-monitoring may have increased treatment implementation for some of the teachers, but self-monitoring plus visual feedback in the form of a graph provided a consistently greater increase in treatment adherence and treatment effect. Still, one month after the visual feedback was withdrawn, treatment integrity was lower or indicated a steady decline (Reinke et al., 2008). A key assertion from these findings is that consistent intervention



implementation requires teachers to change long standing behaviors. Performance feedback must be included to create that change, along with sustained support (Reinke et al., 2008).

It is agreed that heightened teacher treatment integrity is necessary for favorable outcomes for students. The method to monitor integrity can follow indirect or direct methods. Indirect methods appeal to the time and personnel constraints often faced across settings. The problem is that indirect methods, such as self-report, tend to be inaccurate. A great deal of effort is necessary to make self-report a profitable practice. Moreover, it seems that self-monitoring is made effective through consultants providing performance feedback. This may mean at minimum consultants should monitor teacher behaviors and inform the teachers of their adherence to or deviation from the treatment protocol. This demands that consultants continue to monitor and provide feedback as long as the treatment is in place. Ultimately, providing only recommendations fails to guarantee the initial or continued use of an intervention (Noell, 2008).

## **Accountability**

**Role of the Consultant.** Promoting treatment integrity is a demanding task and the question faced by consultants and teachers alike is, “whose responsibility is it?” The notion that treatment integrity is a separate issue from student outcomes is central to answering this question. We are able to identify interventions to improve child behavior, but we are limited in our understanding of how to change adult behavior (Noell, 2008). The field of school psychology has generated research and a framework for consultation that obligates consultants in schools to support intervention implementation; however, all the ways to do so are not as clear (Noell, 2008).

Noell and Witt (1996) examined the assumptions of behavioral consultation and their sentiment was that the process of behavioral consultation has failed to change teacher behavior through solely using interviews and verbal exchanges. First, it was stated that the use of verbal interactions may save time, but they fail to facilitate changes in treatment implementation behaviors. If the consultation process does not result in accurate intervention implementation, then it is not efficient (Noell & Witt, 1996). They also addressed the preference for a collaborative approach to consultation. It is asserted that the habilitative effects of a treatment should carry more weight than teacher preference. Additionally, the field of consultation would benefit the most from detecting consultant interactions that change teacher behavior in respect to treatment integrity (Noell & Witt, 1996). It is assumed that teachers will generalize the skills they have acquired through the consultation process. However, there is little evidence to support this and it is the responsibility to the consultant to program consultation to aid in this endeavor (Noell & Witt, 1996). Lastly, it is thought that the consultant has a very real responsibility to directly assess the consultee and client by using direct measures of student and consultee behaviors to reliably determine intervention effectiveness (Noell & Witt, 1996). The challenge is to develop research to examine the assumptions of behavioral consultation and to have consultants take an active role in assessing the treatments they prescribe.

**Teacher Training.** It is argued that teachers must be held accountable for correct implementation of teaching procedures (Ward, Johnson, & Konukman, 1998). This is one way that school psychologists can begin to assure treatment integrity. The use of feedback and rehearsal contingent upon mistakes was found effective for increasing

specific classroom teaching behaviors across contexts. When given didactic training for teaching behaviors for transitions, lesson introductions, and task presentation, teachers demonstrated the behaviors at low or highly variable rates. After participating in rehearsals for each of the behavior contexts, the teachers performed the behaviors nearly at 100% accuracy across contexts, with little variability in implementation. The rehearsals were employed only if the teachers employed the wrong teaching behaviors (Ward et al., 1998). It was concluded that the use of rehearsal was effective in correcting and making teachers' behaviors consistent across different contexts. Moreover, it was asserted that the avoidance of rehearsals provided the teachers with negative reinforcement for correct implementation. Lastly, the cost to the implement this intervention for teachers is low (Ward et al., 1998). For this set of teachers direct training adequately supported behavioral expectations. Still, their ability to generalize the behaviors over time as the direct training was removed was not examined.

Others have demonstrated differences in implementation of behavioral protocols depending on the type of training. A three phase sequence for training teachers used (1) training through written and verbal information, (2) training with rehearsal, modeling, and performance feedback, and (3) performance feedback (Moore, Edwards, Sterling-Turner, Riley, DuBard, & McGeorge, 2002). The first phase consisted of presenting the teachers with protocols, information and a written test over the intervention procedure. The second phase involved having the teacher practice components of the intervention that were incorrectly implemented and they were given feedback for their overall performance. The third phase required the consultant to only give performance feedback through praise for correct implementation and a verbal review of missteps after each

session (Moore et al., 2002). Percent of correct teacher responses (PCTR) was the primary dependent variable. PCTR in the first phase was variable and low for all the participating teachers. In the second phase, the PCTR increase in level and remained consistently over 95%. In the last phases the PCTR remained as high as the second phase (Moore et al., 2002). This study demonstrated the utility of direct training over didactic training for the use of behavioral protocols. Nevertheless, the researchers did not establish a control condition, which hinders the interpretation of the intervention without any training. These researchers also did not examine if the teacher behavior was maintained with the removal or fading of performance feedback.

One study employed modeling, practice and performance feedback in hopes of training three teachers to write observable and measurable goals, interpret curriculum-based measures (CBM), and calculate progress toward goals and levels of progress. Further, this study included a baseline and maintenance phase (Coddling, Skowron, & Pace, 2005). Teachers were given a CBM training packet with information pertaining to interpretation of data, goal setting, and progress monitoring during the baseline condition. There were no other directions given to the teachers. The modeling, practice and performance feedback condition was the second phase, where teachers were given feedback about their correct and incorrect use of CBM data. Next, the consultant modeled use of the CBM training packet and then had the teacher practice with immediate performance feedback. A new packet was then given to the teachers. The last phase was maintenance, which began a week after the second phase was terminated. It consisted of only providing the teachers with a CBM packet (Coddling et al., 2005b).

At baseline the teachers' implementation of the CBM procedures was 20% to 40% of the steps (Coddington et al., 2005b). All teachers demonstrated either an increase in level or trend in their implementation of the CBM procedures when exposed to the modeling, practice and performance feedback, and they performed about 90% of the steps (Coddington et al., 2005b). During maintenance the results appeared to be mixed. One teacher displayed similar levels of implementation over five observed days and another had similar levels of implementation over three observed days. Yet, one teacher's implementation dropped below 70%, and was given performance feedback, which resulted in her performance trending upward (Coddington et al., 2005b). This study was consistent with the prior studies by demonstrating that modeling, practice and performance feedback is beneficial to teachers' acquisition of skills to follow protocols. Still, due to the inconsistent findings in the maintenance phase and low number of subjects, it is difficult to generalize the authors' optimistic interpretation that the teachers' performance maintained after feedback was removed (Coddington et al., 2005b).

Common across the previous studies is that direct training phases involve the use of modeling, practice and performance feedback simultaneously. One study sought to parcel out the components of this phase, while predicting that more direct training would have a greater impact on teacher treatment integrity. Sixty-four participants were randomly assigned to a didactic training (DT), modeling training (MT), or rehearsal/feedback training (RFT) group (Sterling-Turner, Watson, Wildmon, Watkins, & Little, 2001). DT consisted of verbally explaining the intervention to the teacher. MT participants watched a video tape of an experimenter explaining and demonstrating the intervention, and were given an opportunity to ask questions about the treatment. RFT

participants were given a five-minute training with the researcher and a confederate. Mistakes made by the teacher were corrected immediately and correct implementation was praised (Sterling-Turner et al., 2001). It was found that RFT ( $M = 84.48$ ) yielded a greater percentage of treatment steps adhered to than either DT ( $M = 50.95$ ) or MT ( $M = 70.85$ ; Sterling-Turner et al., 2001). This was a call to provide adequate direct training to consultees so they can implement treatment protocols effectively. These results bolstered previous findings that indicated direct training improves treatment integrity. Moreover, the authors pointed out that without the use of a control group, that generalization is limited (Sterling-Turner et al., 2001).

Continued investigation of the components of teacher training, sought to examine if increasingly direct training of teachers would increase treatment integrity and the effects of the treatment (Sterling-Turner, Watson, & Moore, 2002). Three phases were implemented: (1) baseline for student behavior, (2) didactic training, and (3) direct training. During baseline, no treatment plan was enacted and student behaviors were measured, but teacher behaviors were not. During didactic training the plan was communicated verbally with the consultee and no feedback was given to the consultee. Lastly, implementation data from the first two phases were given to the consultee, praise was given for steps implemented correctly and consistently, and corrective statements with modeling and rehearsal were used for steps performed incorrectly (Sterling-Turner et al., 2002). During baseline, no treatment integrity data was available, since a plan had not been formed yet. After didactic training the level of treatment adherence was low and variable across consultees. Following direct training procedures the level of treatment integrity increased for all consultees and the treatment had a greater effect on decreasing

inappropriate student behavior. Yet, high rates of variability in treatment integrity occurred for some of the participants, even after direct training. This was associated with greater variability in student behaviors (Sterling-Turner et al., 2002). Further, these researchers failed to measure teacher behavior in respect to treatment behaviors at baseline and they did not monitor maintenance after the direct training phase.

**Performance Feedback Emphasis.** The evidence indicates that direct training with performance feedback is the most effective method to increase teacher treatment integrity. The effects of various types of feedback have been examined. One study examined differences in treatment integrity between the follow-up strategies of (1) a weekly meeting, (2) a weekly meeting with commitment emphasis, and (3) weekly meetings with performance feedback. This occurred over the course of three weeks (Noell et al., 2005). The weekly meeting consisted of meeting with and asking the teacher if the intervention was implemented as it was designed. The commitment emphasis meeting was the same as the weekly meeting, except the consultant used a social influence procedure. This procedure included describing how people become distracted from commitments due to constraints, but that follow through is important because of the commitment to the student, potential loss of credibility and harm to the student, evaluating child progress and commitment to the child, and discussed proactive steps to assure treatment implementation. Lastly, the performance feedback meeting consisted of providing positive feedback for steps completed and identifying missed or wrongly implemented steps. Performance feedback was thinned rapidly from daily, to every other day after the teacher reached 100% integrity, to weekly after they reached 100% integrity for two days (Noell et al., 2005). The performance feedback condition was associated

with significantly greater treatment integrity and student outcomes when compared to the weekly and commitment emphasis groups. Moreover, the weekly group and commitment emphasis faced greater deterioration of treatment integrity as time passed (Noell et al., 2005). It appears that the feedback given for a treatment protocol should reflect the teacher's adherence to treatment protocols. Additionally, the feedback should convey if the consultee has adequately or inadequately performed the intervention steps.

Others have questioned if the consultation meeting is necessary and have opted to provide daily performance feedback through the use of graphs, referred to as visual performance feedback (VPF), to show their use of protocol steps (Reinke, Lewis-Palmer, & Martin, 2007). A three day consultation was given to teachers, where the consultant delivered information about the intervention, followed by collection of baseline data. After collecting baseline data for the frequency of teachers' behavior-specific praise, VPF of the previous day's performance was delivered to the teacher prior to instruction each day. After a stable level or trend was identified the teachers entered a maintenance phase. During the maintenance phase data over teacher behavior was taken at 12 and 21 days after cessation of VPF (Reinke et al., 2007). The use of VPF had mixed results, but the teachers generally had an increase in their level of behavior-specific praise. The teachers were found to have lower rates of behavior-specific praise during maintenance, but levels remained higher than at base-line (Reinke et al., 2007). This demonstrates some efficacy for VPF, but due to observed variability in teacher behavior (Reinke et al., 2007) the use of only VPF may be limited. Nevertheless, it may be a valuable component of a more complex system of performance feedback. The collection of maintenance data is limited in its interpretability, since only two points were taken. It shows us that at those moments



some degree of generalization occurred over time, but maintenance took a tertiary role in this study.

### **Maintenance of Treatment Integrity**

**Issues in Performance Feedback Literature.** Early research of teacher training procedures showed some interest in the maintenance of teacher behaviors over time. Carnine and Fink (1978) examined the effects of direct training on the rates of signaling students to perform academic tasks. In this study, multiple baseline and training data points were taken, which strengthened their conclusions that training increases expected teacher behaviors. They measured maintenance five days after the training phase for one teacher and five and twenty days after training for two teachers. It was found that all three teachers presented high levels of the trained behavior at each maintenance observation (Carnine & Fink, 1978), which was interpreted as a high level of maintenance. The problem is that there is limited data to draw conclusions from when comparing phases. It possibly informs us of the level of the behavior, but gives no information regarding variability or trends in teacher instruction implementation during maintenance.

More recent research has included maintenance information at five, ten, and fifteen weeks after performance feedback was removed (Coddling, Feinberg, Dunn, & Pace, 2005a). The researchers collect multiple data points at baseline and performance feedback conditions for the provision of antecedents and consequences for student behaviors. They collected one, two, or three data points for each participant in the maintenance condition through direct observation. Two of the five participants had three maintenance observations, two had two maintenance observations, and one participant

had only one maintenance observation (Coddling et al., 2005a). Similar to Carnine and Fink (1978), all maintenance observations demonstrated high treatment integrity, but there was limited data to adequately draw definitive conclusions. Coddling et al. (2005a) pointed to the possibility that direct observations may induce a reactance in the teachers, which means the observers acted as cues for the teachers to employ the treatment protocol. Lastly, multiple studies provide examples of the inadequacy in the measurement of treatment integrity maintenance (Carnine & Fink, 1978; Coddling et al., 2005a; Coddling et al., 2005b; Reinke et al., 2007; Reinke et al., 2008).

**Revelations in Maintenance Phases.** Some researchers have extended the collection of data in the maintenance phase that provides the field of consultation beneficial information. In a study of the impact of performance feedback on treatment integrity for reinforcement-based treatment, data was collected across training, post-training baseline, performance feedback, and maintenance (Witt, Noell, LaFleur & Mortenson, 1997). The training phases consisted of the teacher practicing the intervention while receiving corrective feedback from the consultant. Post-training and maintenance conditions were identical and consisted of having the teacher implement the intervention independently. The performance feedback condition is when the consultant would provide daily feedback regarding completed and missed steps of the treatment. The range of data points taken for the maintenance condition was from three to eleven (Witt et al., 1997). This allows the reader to examine the effect of performance feedback across time, across participants, and assess performance levels and trends more easily. The data showed that during maintenance treatment integrity was at the same level for each participant. Maintenance data across individual teachers yielded a great deal of variability

with a decrease in level of treatment implementation (Witt et al., 1997). Notably, this study used permanent products to measure integrity and to reduce participant reactance, and this study uses an applied behavior analytic approach to measure maintenance over an extended period of time (Witt et al., 1997).

Research with the administration of educational interventions has followed similar practices as in Witt et al. (1997). Noell, Witt, Gilbertson, Ranier, and Freeland (1997) calculated treatment integrity through the percent of permanent products completed, used an applied behavior analytic approach, and collected extensive maintenance data. The three phases of the experiment were consultation only, performance feedback, and maintenance. The consultation only phase had the consultant develop and explain an intervention to a teacher, and the teacher implemented it independently. Performance feedback involved giving data regarding student performance and teacher treatment implementation, daily. Maintenance was when the teacher implemented the intervention independently and received no contact from the consultant (Noell et al., 1997). Performance feedback produced superior results over the consultation only condition, where treatment integrity was moderate to high. However, examination of the maintenance data revealed that treatment implementation became unstable when performance feedback was removed. In one case, the teacher's treatment integrity decreased enough for performance feedback to be reintroduced (Noell et al., 1997).

Further research of performance feedback and maintenance has looked to fading training and performance feedback. For example, one study compared verbal instruction (VI), faded 3-criteria classroom training (F3CT), response-dependent performance

feedback (RDPF), and maintenance conditions (Gilbertson, Witt, LaFeur-Singletary, & VanDerHeyden, 2007). The teacher was given verbal and written instructions for the intervention in the VI condition and practiced with a role-play scenario. The F3CT was a three step process for reducing the amount of in vivo assistance from the consultant. The first step was delivery of feedback during treatment application until 100% adherence. The second step was delivery of feedback immediately after treatment application until 100% adherence. The third step was delayed delivery of feedback after treatment application until 100% adherence, as determined by examination of permanent products. RDPF consisted of providing positive feedback for implemented intervention steps and a review of the missed steps. The consultant no longer provided performance feedback during the maintenance phase (Gilbertson et al., 2007). They found that in the maintenance condition treatment integrity was variable between and within teachers' application of the intervention. Two of the five teachers' performance dropped significantly enough to warrant the reapplication of RDPF. Further, student performance varied along with the teacher's adherence to the intervention. The authors called for an investigation of variables that would assist in better maintenance of treatment integrity (Gilbertson et al., 2007).

The preceding studies were conducted with one to one performance feedback. In one study, treatment integrity data across performance feedback and maintenance conditions were examined within the context of school-based intervention team meetings. The study consisted of two experiments. The first sought to increase integrity through providing performance feedback after it dropped below a criterion of 70%, where performance feedback would be implemented at the next team meeting. The second

implemented performance feedback at all team meetings, regardless of teacher performance. For the maintenance condition, performance feedback was removed by telling the teachers that they would only have to attend the team meeting if they had any issues with the intervention (Duhon, Mesmer, Gregerson & Witt, 2009). The addition of performance feedback greatly improved teacher treatment integrity in both experiments. The removal of performance feedback was associated with an almost immediate drop in treatment implementation in both experiments (Duhon et al., 2009). Notably, this research demonstrated that providing feedback to teachers at intervention team meetings helps to assure treatment integrity and either improve or maintain integrity (Duhon et al., 2009). This fits with the notion that there needs to be some accountability for intervention implementation and consultants can facilitate that process. Like other studies of performance feedback and integrity, it does not address the challenge of maintaining treatment integrity.

Overall, it appears that shifting from performance feedback to the removal of performance feedback has a detrimental effect on treatment integrity. Performance feedback increases the likelihood that teachers will adhere to intervention steps, the removal of the performance feedback results in decreased treatment integrity. Poorer treatment implementation could mean inconsistent application of the treatment steps, reduced application of the treatment, or failure of the teacher to intervene. Ultimately, this means poorer outcomes for students. It is beneficial to discover what methods can be used to sustain treatment integrity over the course of time, as the consultant withdraws feedback.

**Fading Procedures.** Researchers have attempted to use fading techniques to assure maintenance of treatment integrity. Fading requires adding a cueing stimulus that is gradually reduced in either intensity or components (Zisimopoulos, 2010). In an attempt to find techniques to produce sustained treatment implementation, some have employed fading in tandem with performance feedback, and follow-up meetings were compared to performance feedback. The follow-up meetings required each of the teachers to attend a five minute meeting, where the consultant asked them how the intervention was going and if they had any questions. Performance feedback required the teachers to meet with the consultant and review student academic data, intervention implementation data, and intervention steps completed and missed. When the teacher had four consecutive 100% integrity days, the meetings were changed from daily to every other day (Noell et al., 2000). Three of the five teachers met criteria to have every other day meetings. They continued to demonstrate levels of treatment integrity similar to the daily performance feedback. Mixed results were found for the other two teachers, which the authors took to mean that there needs to be more research to determine effective integrity maintenance procedures (Noell et al., 2000).

Further investigation into fading procedures continued to focus on reducing feedback meetings. Data review procedures were compared to performance feedback with fading procedures (Noell, Duhon, Gatti & Connell, 2002). The data review condition consisted of a daily meeting between the consultant and the teacher, where they reviewed student behavior and the teacher's implementation on the intervention. If the teacher missed steps, they were reviewed and the teacher and consultant then problem-solved to attempt to assure integrity. Sustained integrity resulted in fading the meeting from daily

to every other day (Noell et al., 2002). The performance feedback condition used visual feedback to show the teacher the student's performance and protocol steps followed. When integrity reached 100% for four days, the meetings were reduced from daily to every other day. When they reached 100% integrity for an additional four days, meetings were reduced from every other day to every three days (Noell et al., 2002).

One of the four teachers met criteria for fading in the data review condition. Once fading started, the teacher's implementation of the intervention became highly variable, and her mean adherence went from 100% to 58%. Data review was not effective for any of the other teachers (Noell et al., 2002). Three of the teachers entered the fading process for the performance feedback condition. All of the teachers had a decrease in treatment adherence within four sessions, but rebounded by the third session. However, the great range in percent of steps completed (*range* = 0 to 100%) created too much variability to make any certain conclusions (Noell et al., 2002). The authors concluded that the use of prompts and performance feedback were key to sustaining integrity. This provided accountability for the teachers' adherence to the protocol. The methods they used for fading were closely aligned with typical school-based consultation (Noell et al., 2002).

DiGennaro, Martens, and McIntyre (2005) sought to extend this line of research and hypothesized that the reduction in meetings with a consultant would be negatively reinforcing to teachers. They measured integrity through daily direct observation. They included a performance feedback/negative reinforcement condition and a dynamic fading condition. In the performance feedback/negative reinforcement a consultation meeting was held only after the teachers' integrity dropped below 100%. The meeting consisted of direct rehearsal of the intervention (DiGennaro et al., 2005). They required the teachers to

maintain perfect adherence for three days before they could move to the next phase. The dynamic fading condition entailed giving performance feedback independent of adherence. At first, the meeting occurred every other day, after three consecutive observed sessions of 100% integrity, meetings were reduced to once per week, and thereafter to once every other week (DiGennaro et al., 2005).

The performance feedback/negative reinforcement condition demonstrated an increase in the level of treatment integrity in all of the teachers. Three of the four teachers were able to attain the 100% criterion; however, the criterion for one teacher was negotiated to 60%, based upon the teacher's stated unwillingness to carry out the intervention. Further, this teacher's adherence was highly variable during this phase (DiGennaro et al., 2005). The dynamic fading procedure demonstrated treatment integrity levels across consultees similar to the prior condition across participants. Still, there was a great range in the actual implementation across participants. Only one teacher maintained 100% integrity, while two showed variability of 20% or more, and one ranged from 70 to 80% total integrity (DiGennaro et al., 2005). It was conjectured that the processes of performance feedback/negative reinforcement and dynamic fading increased treatment integrity by providing adequate support for teachers implementing interventions (DiGennaro et al., 2005). Important limitations to this study were highlighted and these included the use of direct observation, delays in reinforcement using a negative reinforcement model, and failure to inform the teachers immediately that they failed to implement the treatment correctly. Direct observation may have cued teachers to perform the intervention (DiGennaro et al., 2005) and direct observation is often impractical in naturalistic school settings. The application of negative



reinforcement could delay reinforcement and cause a decrease in performance of the treatment steps (DiGennaro et al., 2005). Lastly, failure to alert teachers of poor adherence may reinforce behaviors incompatible with adherence to the protocol (DiGennaro et al., 2005).

A following study built upon the notions that reinforcement is key to the integrity of treatment implementation and that fading procedures can maintain treatment integrity. To examine this, the researchers used a multiple baseline design across subject with six phases: (1) pretraining baseline, (2) training, (3) implementation baseline, (4) goal setting and student performance feedback, (5) teacher performance feedback and direct rehearsal with meeting cancellation, and (6) fading. Four teacher-student dyads participated and integrity was measured through direct observation (DiGennaro, Martens & Kleinmann, 2007).

Pretraining baseline occurred when the researchers observed the teachers' behaviors before the treatment was introduced. Training included direct instruction, modeling, coaching, and corrective feedback for the intervention and it continued until all steps were demonstrated at 100% integrity on two occasions with the consultant's assistance. Teachers implemented the intervention without feedback or assistance during implementation baseline (DiGennaro et al., 2007). Goal setting and student performance feedback was devised to improve integrity by showing the teachers goals for their respective student and the students' progress. The teachers were provided daily graphed feedback for their treatment integrity. If they had 100% integrity, then the consultant meeting was cancelled; if integrity was below 100%, then they had to meet with the consultant for teacher performance feedback and direct rehearsal with meeting

cancellation. Lastly, fading was designed to reduce meetings to every other day, once per week, and once every other week for every three days at 100% integrity (DiGennaro et al., 2007).

As expected, none of the teachers demonstrated behaviors consistent with the steps of the protocol during pretraining baseline. During training there large variability between teachers, for example, one teacher implemented it at 100% integrity for two sessions immediately after pretraining, and one teacher took 15 sessions and ranged from 0 to 100% integrity (DiGennaro et al., 2007). At implementation baseline, all teachers dropped significantly in treatment adherence; three plummeted to 0% adherence in three sessions or less, and one trended downward, reaching as low as 30% adherence (DiGennaro et al., 2007). Goal setting and student performance feedback demonstrated mixed results across participants, but all had low integrity. Three of the teachers reached 100% integrity after four sessions or less with performance feedback and direct rehearsal with meeting cancellation, while one teacher took seven sessions (DiGennaro et al., 2007). During the fading procedure only one of the three teachers received the full fading procedure because of time constraints. She maintained integrity at 90% or better for 13 observed sessions (DiGennaro et al., 2007). The other teachers maintained an average adherence above 90% through the first portion of fading, and one teacher did so through two fading procedures (DiGennaro et al., 2007).

The researchers concluded that direct training with performance feedback was the most effective method for increasing teacher treatment integrity. The component of meeting avoidance was part of the performance feedback and felt to provide a means negative reinforcement. Nonetheless, it was not discerned what aspects of this condition

had the greatest effect on integrity (DiGennaro et al., 2007). The researchers mentioned that the use of fading for treatment integrity is an area that needs more research (DiGennaro et al., 2007). They may have found it more fruitful to investigate direct training with performance feedback and fading procedures exclusively. The use of five phases, exclusive of the fading procedures may have led them to their time constraints. Moreover, DiGennaro et al. (2005) mentions that the repeated use of trials that result in poor adherence could reinforce failure in following the protocol. It would have been more efficacious to demonstrate the usefulness of training, implementation baseline, direct training with performance feedback, and then fading. Time is saved because of the elimination of two phases that have demonstrated an association with poor integrity: (1) pretraining baseline, and (2) goal setting and student performance feedback. The steps mentioned above seem more feasible and practical for supporting teacher accountability for treatment implementation. Further, additional components should be integrated into this structure to enhance the effects of generalizing treatment integrity over time.

### **Indiscriminable Contingencies**

**Indiscriminable Contingencies.** The use of indiscriminable contingencies is a practical strategy to support behavior maintenance (Freeland & Noell, 2002). Stokes and Baer (1977) pointed out that behavior change must happen over time, settings and affect related behaviors. In order to accomplish this they suggest that generalization should be programmed into procedures and not just assumed as a byproduct of intervention. Generalization is then best conceptualized as the occurrence of a relevant behavior under conditions different than scheduled training conditions (Stokes & Baer, 1977). One method to reach this end is indiscriminable contingencies. Indiscriminable contingencies

rely on the use of intermittent schedules of reinforcement to create a situation where the individual does not discriminate a time of reinforcement from a time of non-reinforcement, and thus continues to emit the behavior over time. This creates behaviors that are highly resistant to extinction, as well (Stokes & Baer, 1977). Additionally, indiscriminable reinforcement contingencies create a state where the individual will respond to the presence of the stimulus without actual reinforcement delivery, is advantageous over fixed schedules of reinforcement for generalization, and sustains treatment effects over time better than continuous reinforcement (Stokes & Baer, 1977).

**Examples in Student Intervention.** Some studies have demonstrated the usefulness of indiscriminable contingencies for interventions with children. One study paired the use of continuous reinforcement and indiscriminable contingencies to increase prompted behaviors (Guevremont, Osnes, & Stokes, 1986). The indiscriminable contingencies phase consisted of delivering a mixed sequence of contingencies. The sequence first consisted of continuous delivery of reinforcement, followed by reinforcement for observed desired behaviors, but not solicited desired behaviors. Then there was no delivery of reinforcers, preceded by delayed reinforcement, after which there was reinforcement of some desired behaviors, rather than others. Finally, some days there was no reinforcement for any desired behaviors (Guevremont et al., 1986). It was found that the sequence used for indiscriminable contingencies was successful in maintaining behaviors over time. This was interpreted as the sequential change in the contingencies helped to maintain the behavior because the participants could no longer discriminate between reinforcement contingencies, whereas abrupt stops in reinforcement signal the end of the reinforcement delivery (Guevremont et al., 1986).

Others have supported the assertion that the use of intermittent reinforcement schedules makes the fading of a reinforcement contingency indiscriminable. When intermittent reinforcement is combined with a delay in reinforcement, it becomes difficult for individuals to differentiate between the shift from one treatment phase to the next (Freeland & Noell, 2002). The combined use of delay and intermittent scheduling of reinforcement was examined as a method to create indiscriminable contingencies for math skill performance. This was examined through baseline, one reinforcement phase, two delay phases, and one maintenance phase (Freeland & Noell, 2002). Baseline was the assessment of digits correct per minute from a math worksheet, and the child was given no incentives or feedback. A goal was set for digits correct per minute and when it was surpassed the student received an incentive from a goody box during the reinforcement stage. The first delay condition was Delay-2, and after two worksheets were completed, the researcher randomly chose one to be graded. If it met or surpassed the goal, then the child received a reward. The second delay condition, called Delay-4, had the researcher randomly choose one math worksheet after four were completed. Again, meeting or surpassing the goal resulted in reward. The maintenance phase consisted of providing the goal to the child by writing it on the worksheet; otherwise, it was the same as baseline (Freeland & Noell, 2002).

It was found that the children responded to the reinforcement conditions with increased levels of performance in digits correct per minute when compared to baseline. When the Delay-2 and Delay-4 conditions were applied, their performance was consistent with the reinforcement condition. When the maintenance phase was implemented, the students had an increase in their level of performance, one student for 18 days before it

became variable and the other for 24, when the study was discontinued (Freeland & Noell, 2002). The authors concluded that delayed and intermittent reinforcement contributes to the maintenance of responding over time. Maintenance occurred as reinforcement thinned and there was an increase in response effort to obtain the reinforcer (Freeland & Noell, 2002). This informs us that the use of reinforcement may be necessary to establish a behavior. After the individual demonstrates the behavior in an accurate and fluent manner, adjusting reinforcement to randomized contingencies while systematically creating delays could promote maintenance of desired behaviors.

**Application to Teacher Treatment Integrity.** Studies have shown poor maintenance of teacher treatment integrity (Duhon et al., 2009; Gilbertson et al., 2007; Noell et al., 1997; Witt et al., 1997). In response, some have attempted to apply fading procedures to maintain teacher treatment integrity (DiGennaro et al., 2005; DiGennaro et al., 2007; Noell et al., 2000; Noell et al., 2002). These studies have ignored is the use of indiscriminable contingencies within consultation to increase the likelihood of treatment integrity maintenance. Also, these studies have failed to systematize the fading procedures into distinct phases, as done in Freeland and Noell's (2002) study of math fluency.

It seems sensible to include indiscriminable contingencies when programming maintenance of behaviors in treatment integrity, given its efficacy in maintaining other behaviors. Most interesting is the use of delayed, random selection of a work product to examine adherence. This could be used with byproducts of treatment implementation, such as permanent products. Delaying the drawing of the permanent product over the course of one day, one week, and then two weeks would merge with the previously used

fading procedures (DiGennaro et al., 2005; DiGennaro et al., 2007) and indiscriminable contingencies tactics (Freeland & Noell, 2002; Guevremont et al., 1986). Ultimately, the reduction to the two week period seems arbitrary. To be more relevant, studies of fading procedures should fade to a regularly occurring event, which includes the appraisal of treatment application to student problems.

### **Prereferral Intervention Teams**

In school settings, prereferral intervention teams meet at regularly scheduled times during each month of the school year. Prereferral intervention teams are multidisciplinary teams that meet to engage in problem-solving to increase teacher support and enhance student outcomes. These go by many names, such as the Student Assistance Team, Intervention Assistance Team, Child Study Team, Instructional Support Team (Burns, Vanderwood, & Ruby, 2005). For simplicity, from here onward they will be referred to as Student Assistance Team (SAT). SATs are felt to be one of the most effective means to resolve problem behaviors in children and are liked for their school-based and problem solving nature (McDougal, Nastasi, & Chafouleas, 2005). Moreover, 86% of the states within the United States recommend or mandate their use and more are likely to follow (Truscott, Cohen, Sams, Sanborn, & Frank, 2005). SATs are developed as an alternative to the traditional model of refer-test-place for special education (McNamara, Rasheed, & DeLamatre, 2008). They follow a general individual problem-solving model of referral, problem analysis, problem statement, goal setting, intervention development, support of intervention implementation, progress monitoring, and outcomes evaluation (Bahr & Kavaleski, 2006; Burns et al., 2005; McNamara et

al.,2008). The emphasis on problem-solving helps to decrease the sentiment that SATs are only a stepping stone to special education placement (Bahr & Kavaleski, 2006).

The student brought to the SAT process often have severe problems, which warrant sound intervention and progress monitoring (Truscott et al., 2005). SATs effectively develop interventions to help children in need within the general education setting without placement or testing to determine disability (McDougal et al., 2005). A large portion of the teams' job is to know effective interventions for specific problems, and to collect and review data to determine severity of the problem and effectiveness of treatment within the classroom (Bahr & Kavaleski, 2006). As part of the SAT process teachers help identify problems, develop interventions, gather data, and make decisions; however, it is the responsibility of all members to have a role in the approach (Slonski-Fowler & Truscott, 2004; Truscott et al., 2005). Members, such as school psychologist, can provide regular support to the teacher through integrity evaluations because of SAT's consistent and on-going nature (McDougal et al., 2005; Bahr & Kavaleski, 2006).

SAT members could help teachers increase treatment adherence by giving regular feedback at the initial implementation of an intervention, and by giving periodic performance feedback as the treatment is underway. The team can further assist teachers by helping them to understand the rationale of the intervention in order to secure implementation (McDougal et al., 2005). This prevents teachers from falling back on mass market curriculum packages, which may have questionable empirical foundations, to intervene on the student (Bahr & Kavaleski, 2006). More importantly, these activities within the team help to validate teacher input, fit interventions with problems, and hold teachers accountable for their actions. These points are found to increase teacher



involvement by making them feel valued, providing roles for those involved, and communicating expectations clearly (Slonski-Fowler & Truscott, 2004).

### **Multiple Baseline Designs**

Multiple baseline across subjects research designs have been used in studies that examine intensity of consultant involvement (Hagermoser-Sanetti & Kratochwill, 2009; Reinke et al., 2008), the use of direct training and performance feedback to increase treatment integrity (Coddling et al., 2005b; Moore et al., 2002; Reinke et al., 2007; Sterling-Turner et al., 2002; Ward et al., 1998), various methods to support acquisition and demonstrate maintenance of treatment implementation (Carnine & Fink, 1978; Coddling et al., 2005a; DiGennaro et al., 2005; DiGennaro et al., 2007; Duhon et al., 2009; Gilbertson et al., 2007; Noell et al., 1997; Noell et al., 2000; Noell et al., 2002; Witt et al., 1997), and maintenance in response to indiscriminable contingencies (Freeland & Noell, 2002; Gueveremont et al., 1986). This design is well received because it addresses socially relevant behaviors in their natural setting and it does not require the withdrawal of an intervention to show a treatment effect. Single subject designs that require intervention withdrawal to show treatment effects may be problematic when some behaviors, such as academic skills, are not reversible, and other behaviors may be unethical to reverse, such as cessation of self-injury (Baer, Wolf, & Risely, 1968; Tawney & Gast, 1984). Multiple baseline designs protect against history, maturation, and sequencing effects by extending the baseline of all the participants (Carr, 2005; Noell & Gresham, 2001). These designs are useful when a detailed single-subject design is needed to demonstrate a treatment effect over time (Noell & Gresham, 2001).

The researcher, in using multiple baseline research, attempts to demonstrate a reliable outcome for an intervention by showing the greatest change in behavior only under the application of the treatment (Baer et al., 1968). Experimental control is demonstrated through baseline data across three or more conditions at the same time. When all three demonstrate stability in the level and trend of the behavior, the intervention is given in the first condition. There is usually an abrupt change associated with the intervention, while the other baselines remain the same. When a specified level of performance is reached in the first condition, the treatment is given in the second condition, and so on (Tawney & Gast, 1984). In general, all forms of multiple baseline designs follow the same guidelines: (1) decide upon outcome objectives before beginning the study, (2) intervene only after all baselines are stable, (3) intervene only after the participant has reached the criterion set before the study, (4) have three or more baselines, (5) identify functionally independent baselines to avoid behavioral covariation, and (6) use similar baselines to evade inconsistent intervention effects (Tawney & Gast, 1984). Conclusions are drawn by directly visually inspecting the data, and this design aims to show replication of results across subjects (Baer et al., 1968). Multiple baseline design uses three elements of single-case design logic. The first is repeated measures can establish the prediction of a baseline's path, which will allow for the detection of differences between the baseline and treatment phase path. Second, the effects of the intervention are verified by inferring that changes in one participant's behaviors were independent of other participants' behaviors. Third, the impact of the treatment is replicated across individuals (Carr, 2005).

Specific to multiple baselines across subject designs is the investigation of an intervention across three or more individuals, who emit similar levels of a target behavior; this type of design is intersubject. Other multiple baseline designs look at the same subject across behaviors or settings; these designs are intrasubject (Tawney & Gast, 1984). The guidelines are similar to those for multiple baseline designs in general. They are: (1) decide upon outcome objectives before the beginning the study, (2) intervene only after all participants demonstrate stable patterns of behavior, (3) intervene only after the preceding participant has reached the criterion, (4) have three or more participants, (5) identify functionally independent subjects, and (6) use similar baselines subjects (Tawney & Gast, 1984). One advantage to this design is that it targets skills needed across individuals. Other strengths include staggering time across individuals to allow for different rates of learning and shows effectiveness across participants, which increase external validity (Tawney & Gast, 1984). There are two problems with the across subjects design. First, keeping individuals independent of each other, so that baselines stay stable until treatment is applied. Assuring the subjects are similar enough to experience the same effects as a result of the treatment is the second (Tawney & Gast, 1984). This design is limited because it requires three individuals with the same history and functioning at baseline. Another drawback is that baseline data that are collected concurrently and continuously may be impractical in applied settings. Lastly, it may be questionable to postpone intervention when a participant is ready to learn it, especially if it will benefit him or her (Tawney & Gast, 1984).

The data for multiple baseline designs are often placed into a graphic representation in the form of a line graph. This is where consecutive data points are

marked on a two dimensional plane and connected with a line, called the data path.

Measures of performance typically correspond to placement on the vertical axis, whereas the horizontal axis represents the passage of time (Cooper, Heron, & Heward, 2007).

Visual analysis is used to determine if data points show variability, a particular level of behavior, and any trends in behavior. Variability is the difference in outcome along multiple measures of the same behavior. An increase in variability would look like an exaggerated vertical zigzag. Less variability would appear more linear (Cooper et al., 2007). The level of a behavior is its central tendency (i.e. mean or median) on the vertical axis. A vertical line is drawn through the array of data points to demonstrate its level (Cooper et al., 2007). Trend is the overall direction that a data path takes and is described in direction (increase or decrease), degree (gradual or steep), and variability of the data points around the trend (Cooper et al., 2007).

## CHAPTER III

### METHODOLOGY

#### **Overview of the Student Assistance Team**

The SAT was a multidisciplinary, pre-referral team assembled to identify areas of concern for children, develop interventions, assess the effectiveness of those interventions, and make recommendations for extended support based upon their findings. These were typically comprised of a school principal, school psychologist, the referred student's teacher, an education specialist, and other parties of concern. Interventions developed through the SAT process focused on either increasing academic skills or altering social behaviors of students. The SAT meeting required members to examine graphic or written data regarding student progress and receive feedback about intervention implementation. Ideally, the feedback for intervention implementation included measures of treatment integrity, such as adherence to the intervention's steps.

#### **Overview of Procedures**

Treatment protocols were developed to address specific referral concerns brought through the SAT process. Each treatment protocol had a resulting permanent product, such as marked areas of an intervention sheet and sticker charts. The teachers were trained to implement the respective treatments for their students and collect the

permanent products. The teachers were responsible for completing the treatment protocols and placing the permanent products in the intervention folders. The primary investigator checked the products from the intervention folder daily and gave scheduled performance feedback to the teachers, and collected the products at the conclusion of the study. When a student was absent, the materials were not checked for that day, and the intervention was continued upon the child's return.

### **Participants**

Participants were selected from three different schools located within a mixed suburban/rural school district in the South-Central United States. Three teacher and student dyads were selected to participate in the study. The dyads were referred to the SAT for help with academic or behavioral problems and the primary investigator acted as a consultant to the teachers. The procedures for the study were explained to the teachers and incorporated as part of the SAT process for these teachers. Informed consent for participation in the research project was obtained from both the participating teacher and the parents of children receiving services as part of the study. The initial consultation sessions occurred during the last week of January. All of the participating teachers were Caucasian women and all of the participating students were girls. All of the teachers reported that they never experienced a SAT process were they received structured feedback regarding treatment integrity. The names of the teachers and students were changed to unrelated names to protect their anonymity.

Ms. Chief and Agnes comprised the first dyad. Ms. Chief was a first-grade teacher with 17 years of experience and she held state certification in early childhood education. She referred Agnes to the SAT process for poor work completion, aggressive behaviors,

and disruptions to the classroom. Agnes' behaviors included throwing objects and desks around the classroom, hitting peers and school faculty, and swearing at school staff. Ms. Sal and Agatha comprised the second dyad. Ms. Sal was a second-grade teacher with eight years of experience and she held state certification in elementary childhood education. She referred Agatha to the SAT process for defiance towards school personnel, and aggression and intrusion towards peers. Agatha's behaviors included mocking school staff, roaming around the classroom and the school, and running from staff when she was confronted. She intruded into peers' space during seat work and had a history of threatening peers and running into them in the hallway. Ms. Tiffye and Jacqe comprised the third dyad. Ms. Tiffye was a kindergarten teacher with one year of experience and she held state certification in early childhood education. She referred Jacqe to the SAT process for deficits in letter naming skills. Jacqe reportedly could identify less than half of the upper and lower case letters in the English alphabet accurately.

### **Setting**

The problem assessment used to develop the interventions occurred within each teacher's respective classroom. All initial training and performance feedback sessions took place in the teacher's classroom. Once feedback was thinned to every other week, performance feedback was delivered as part of the SAT meetings. SAT meetings occurred every two weeks in a conference room. The teachers carried out the interventions within their respective classrooms.

## **Materials**

**Problem assessment measures.** Agnes and Agatha were observed in their respective classrooms using partial interval sampling observation forms at 12 second intervals for 10 minutes each observation during a period, which their respective teachers reported high amounts of misbehavior. Agnes was on-task 44% of the observed intervals and her behavior consisted of attempts to engage Ms. Chief in conversation and walking around the room to gain Ms. Chief's attention. To determine the impact of reward with feedback on Agnes' behavior, she was offered a sticker for sitting in her chair, demonstrating listening behaviors, and completing her seat work for three 10 minute intervals during the morning work period. At the end of the three 10 minute periods she was given feedback about her performance and she was on-task an average of 95% of the time when she could earn a sticker. She had a marked decrease in attention seeking behaviors and the amount of attention she received for misbehavior. Additionally, Ms. Chief kept an antecedent-behavior-consequence log for two weeks prior to the implementation of the intervention, which showed Agnes averaged one instance of throwing desks, objects, hitting others, and swearing at school faculty daily.

Agatha was on-task for 50% of the observed intervals and her behaviors consisted of walking around the room to engage others in conversations, and she verbally and physically intruded upon peers. Ms. Sal employed a token economy where students earned "dollars" throughout the school day and at the end of the week they trade in the "dollars" to participate in fun activities. Agatha was offered five extra dollars each time she could stay in her tape-marked seat area, work quietly, and complete a segment of her assignment for 10 minutes. Agatha was on-task an average of 75% of the time when she



could earn extra dollars over three 10 minute periods. Her verbal behavior and movement around the room trended downward across all trials; her levels of listening to the teacher, completing seat work, and participating all increased in level.

Jacqe's letter naming fluency (LNF) was assessed with three LNF probes from the DIBELS® Kindergarten Benchmark Assessment. She had a median of 8 letters named correctly in one minute, which placed her in the 4<sup>th</sup> percentile for kindergarteners at the winter benchmark. Her letter naming accuracy was assessed by presenting her with a list of upper case letters and then a list of lower case letters, and having her name as many as she could. The letters were pointed to at random. She correctly named 13 of 26 upper case letters and 11 of 26 lower case letters, or named 24 of 52 (46%) letters correctly. A skill/performance deficit assessment is an assessment designed to determine whether a student has the skills necessary to perform a given task or if low performance is a result of a lack of necessary skill. This is measured by offering the student a reward for improved performance on a previous task. (Duhon et al, 2004). A skill/performance assessment was conducted with Jacqe and it yielded no difference in her performance for letters named correctly. She was determined to have a skill deficit for letter identification.

**Intervention materials.** Materials for this study included printed intervention protocols, folders, daily check sheets, as well as other materials required to complete the individualized interventions. Printed protocol were created based upon the necessary treatment steps of a given intervention, which details each step of the treatment and how to create and collect permanent products. Each teacher received a folder for the collection of permanent products along with daily check sheets of the treatment steps, and all other materials necessary to complete the intervention.

A check-in/check-out type intervention was developed with Ms. Chief to increase Agnes' appropriate behaviors in the classroom. The intervention employed procedures to provide feedback after 10 minutes of instruction during each morning work period and allows for Agnes to earn incentives when demonstrating specific levels of appropriate classroom behaviors. The intervention included an intervention protocol form that stated materials needed, instructions, and operationally defined expectations for Agnes (Figure 1). The check sheet was identical to the protocol sheet, except it had short blanks in front of each number. Daily score recording forms were issued that had an area to write the date, and a grid that allowed for the teacher to rate each expectation at each time period and total the points Agnes earned. Under the grid was a guide for what Agnes earned depending upon the total number of points for the morning work period and cues for feedback for each of the expectations (Figure 2). Ms. Chief received a sticker chart for Agnes to place stickers on, which indicated how close she was to earning a prize (Figure 3). Agnes' prize consisted of earning extra time to read with older students after earning eight stickers.

A check-in/check-out type intervention was developed with Ms. Sal to increase Agatha's appropriate behaviors in the classroom. The intervention employed procedures to provide feedback after specific instructional periods during each morning and allows for Agatha to earn incentives when demonstrating specific levels of appropriate classroom behaviors. The intervention included an intervention protocol form that stated materials needed, instructions, and operationally defined expectations for Agatha (Figure 4). The check sheet was identical to the protocol sheet, except it had short blanks in front of each number. Daily score recording forms were issued that had an area to write the

date, and a grid that allowed for the teacher to rate each expectation at each time period and total the points Agatha earned. Under the grid was a guide for what Agatha earned depending upon the total number of points for the morning work period and cues for feedback for each of the expectations (Figure 5). Ms. Sal received the same sticker chart for Agatha that Ms. Chief received for Agnes. Agatha's prize consisted of receiving 10 extra classroom "dollars" after earning eight stickers.

A letter-identification intervention with error correction and reward was developed with Ms. Tiffye to increase Jacqe's letter identification accuracy. This was designed to target upper and lower case letters and provide incentive for Jacqe to increase letters she named correctly. The intervention included an intervention protocol form that stated materials needed and instructions (Figure 6). The check sheet was identical to the protocol sheet, except it had short blanks in front of each number. A daily recording sheet was provided for upper (Figure 7) and lower (Figure 8) case letters, and included grids for multiple days on each sheet. Ms. Tiffye received the same sticker chart for Jacqe that Ms. Chief and Ms. Sal received for their students. Jacqe's prize consisted of choosing a small toy after earning eight stickers.

### **Dependent Variables**

**Treatment Integrity.** The primary dependent variable for the study was the level of treatment integrity that the teachers exhibited. For this study treatment integrity consisted of the percentage of treatment steps completed as part of the treatment protocol. A completed treatment step was measured by evaluating permanent products. Treatment integrity was computed by dividing the number of completed treatment steps by the total number of expected treatment steps, multiplied by 100%. Ms. Chief and Ms. Tiffye had 12 steps, and Ms. Sal had 13 steps in the intervention protocol given to them.

**Student Outcomes.** Student data generated from the implementation of the treatments was collected as student outcome data. This data were specific to either the academic or social behavior concern targeted and obtained from the permanent products produced by implementation. Throughout the study, academic and social behavior student outcome data were collected and reported but was not evaluated within the experimental design. Agnes and Agatha had their classroom conduct rated by their respective teachers using the point system from the intervention protocol; Agnes could earn up to 18 points and Agatha could earn up to 30 points. The number of letters named correctly each day, at the beginning of the intervention served as daily progress monitoring for Jacqe.

### **Experimental Design**

A non-concurrent multiple baseline design across subjects was used to evaluate the effects of the consultation procedures on teacher treatment integrity. The procedures consisted of a direct training phase, an implementation baseline, performance feedback with direct rehearsal, three fading phases, and a maintenance phase. The direct training phase was designed to provide the teacher an opportunity to acquire adequate skills necessary to perform the intervention. The purpose of the implementation baseline was to examine the level of teacher treatment integrity independent of external support. Performance feedback with direct rehearsal was employed in the event that low levels of treatment integrity were observed in an attempt to increase treatment integrity. The fading phases were designed to systematically decrease the amount of performance feedback with direct rehearsal required to maintain high levels of treatment integrity. The

maintenance phase attempted to reduce feedback to that which occurs naturally as part of the SAT process.

**Direct Training Procedures.** Soon after the initial SAT referral meeting, each teacher met individually with the consultant to read and review the treatment protocol designed for the targeted concern, and allow for any clarification or alterations to the plan. The consultant read and demonstrated each step in the treatment protocol to the teacher. The consultant demonstrated how to make and collect the permanent products. Then the teachers read and demonstrated each treatment component as well as completed practice permanent products in a role play scenario. In the role play the teacher administered the treatment without any assistance from the consultant who acted as the student. If teacher treatment integrity was below 100% they received corrective feedback during the role play and the role play was repeated. Once the teacher has reached 100% integrity without assistance, the training ended and the teacher was given copies of the training materials as references for delivery of the treatment and completion of permanent products.

**Implementation Baseline.** The teachers were given copies of the materials from the direct training session during the implementation baseline (IB). The teachers also received all material necessary to complete the intervention independently, intervention folder, and procedural check sheets. During this phase the teachers were asked to complete the intervention and received no other contact from the consultant, except for daily collection of the intervention folder. The teachers were informed that at the next SAT meeting, the team and teacher will: (1) receive a graph representing adherence to treatment steps, (2) review the percentage of intervention steps implemented and (3)

discuss the student's progress. The implementation baselines lasted three days for Ms. Chief, eight days for Ms. Sal, and 17 days for Ms. Tiffye. This was to evaluate the effect of the SAT feedback procedures on integrity prior to any more intensive procedures.

**Performance Feedback with Direct Rehearsal.** If after the IB SAT meeting occurred and teacher treatment integrity was low, highly variable or trending downward, then the teachers received performance feedback with direct rehearsal (PFDR). Low integrity was adherence to treatment steps averaging below 80% or trending downward and reaching below 80% during IB. Prior research has identified high treatment fidelity as at least 80% (Borelli et al., 2005). Integrity was considered too variable if adherence to treatment steps had a range of more than 20%. Downward trends in treatment integrity were assessed through the visual inspection of the graphed percent of treatment steps followed daily. During this phase the teacher received daily: (1) a graph demonstrating the percentage of permanent products completed during IB through PFDR, (2) feedback regarding the missed steps, and (3) rehearsal of the entire treatment protocol until 100% adherence. PFDR continued each day until the teacher demonstrates 100% integrity for two prior consecutive days.

**Fading 1.** Once teacher treatment integrity reached 100% for two preceding days during the PFDR phase, then the fading 1 schedule (F1) began. The teacher was informed that integrity will be determined by randomly choosing one day's permanent products. The consultant and teacher's daily meetings consisted of: (1) a graph demonstrating the percentage of permanent products completed during the IB, PFDR, and F1 phases, (2) praise only, if there was 100% completion of permanent products, or performance feedback with rehearsal if there was less than 100% completion of permanent products

and (3) a review of the student's progress. The teacher meeting occurred in the teacher's classroom.

**Fading 2.** If teacher treatment integrity reached 100% for two preceding days during the F1 phase, then the teacher began with the fading 2 schedule (F2). During F2 the teacher was informed that there will only be meetings every other day. The teacher was informed that integrity will be determined by randomly choosing one day's permanent products. The consultant and teacher meetings consisted of: (1) a graph demonstrating the percentage of permanent products completed during the IB, PFDR, F1 and F2 phases, (2) praise only, if there was 100% completion of permanent products, or performance feedback with rehearsal if there was less than 100% completion of permanent products and (3) a review of the student's progress. The teacher meeting occurred in the teacher's classroom.

**Fading 3.** If teacher treatment integrity reached 100% for the two previous days during the F2 phase, then the teacher began with the fading 3 schedule (F3). During F3 the teacher was informed that there will only be one weekly meeting, which will correspond with the same day of the week as the SAT meeting. The teacher was informed that integrity will be determined by randomly choosing one day's permanent products. The consultant and teacher meetings consisted of: (1) a graph demonstrating the percentage of permanent products completed during the IB, PFDR, F1, F2 and F3 phases, (2) praise only, if there was 100% completion of permanent products, or performance feedback with rehearsal if there was less than 100% completion of permanent products and (3) a review of the student's progress. The teacher meeting occurred in the teacher's classroom.

**Maintenance.** If teacher treatment integrity reached 100% for the two previous days during the F3 phase, then the teacher began the maintenance phase (MA). During MA the teacher was informed that she will only have to meet one time every other week for the SAT meeting. The teacher was informed that integrity will be determined by randomly choosing one day's permanent products. Consultation at the SAT meeting consisted of: (1) a graph demonstrating the percentage of permanent products completed during the IB, PFDR, F1, F2, F3 and MA phases, (2) praise only, if there was 100% completion of permanent products, or performance feedback regarding the steps that were not followed if integrity was below 100%, and (3) discussion of the student's progress.



## CHAPTER IV

### FINDINGS

#### **Treatment Integrity**

The primary purpose of this study was to examine teacher treatment integrity. The teachers experienced seven phase changes: direct training (DT), implementation baseline (IB), performance feedback with direct rehearsal (PFDR), fading 1 (F1), fading 2 (F2), fading 3 (F3), and maintenance (MA). All three teachers' performance in percent of steps completed correctly was examined through visual analysis of the graphs generated from the data (Figure 9). The IB and MA means and standard deviations were taken for each teacher, and used to calculate Cohen's  $d$  (Table 1) to determine standardized performance differences. All three teachers demonstrated 100% integrity during DT, low treatment integrity at IB, and improved performance through the PFDR, F1, F2, F3, and MA phases.

Ms. Chief's IB data were collected over three consecutive days and MA data were collected over 10 opportunities to implement the intervention. Her IB did not contain a SAT meeting because Agnes' behaviors were too extreme to be allowed to continue without adequately implemented intervention (i.e., physically attacked school staff, threw books and desks around the classroom, swearing and threats directed at school staff and peers). Ms. Chief completed a mean of 75% ( $SD = 0$ ) of the intervention steps during IB and completed a mean of 88% ( $SD = 31.48$ ) of the intervention steps during MA. She had a medium improvement from IB to MA (Cohen's  $d = 0.60$ ). Her performance at IB was consistent and low. She completed the PFDR, F1, F2, and F3 phases in 10 days with low variability, where she missed one step on one day. Ms. Chief's performance during MA was at a high level, but had some variability. There were two days where she completed less than 100% of the intervention steps, where one day she was still in the acceptable range. The one data point she was at 0% treatment integrity appears to be an outlying event.

Ms. Sal's IB data were collected over seven consecutive days and MA data were collected over eight opportunities to implement the intervention. Ms. Sal completed a mean of 54% ( $SD = 14.8$ ) of the intervention steps during IB and completed a mean of 99% ( $SD = 2.83$ ) of the intervention steps during MA. She had a large improvement from IB to MA (Cohen's  $d = 4.26$ ). Her performance at baseline was low and variable, and she demonstrated a slight upward trend following the IB SAT meeting, which leveled off after a short period. She completed the PFDR, F1, F2, and F3 phases in 13 days with variability, where she missed multiple steps on two days during PFDR and F1, but missed none in the F2 and F3 phases. Ms. Sal's performance during MA was at a high level with

minor variability. She completed less than 100% of the intervention steps for one day, but her performance was still at a high level.

Ms. Tiffye's IB data were collected over 14 opportunities to implement the intervention and MA data were collected over 12 opportunities to implement the intervention. Ms. Tiffye completed a mean of 68% ( $SD = 32.66$ ) of the intervention steps during IB and completed a mean of 98% ( $SD = 3.62$ ) of the intervention steps during MA. She had a large improvement from IB to MA (Cohen's  $d = 1.3$ ). Her performance at baseline was low and highly variable before and after the IB SAT meeting. She completed the PFDR, F1, F2, and F3 phases in eight days with consistently high performance across all phases. Ms. Tiffye's performance during MA was at a high level with minor variability. She completed less than 100% of the intervention steps for three days, but her performance was still at a high level.

### **Comparison of Treatment Integrity and Student Performance**

A secondary purpose of this study was to examine the impact of treatment integrity on student performance. This was done through examining the performance of the students across IB and MA phases and then comparing teacher treatment integrity against their respective student outcomes. Student outcomes were examined using each student's respective performance measure. Teacher treatment integrity and student performance comparisons were made using respective teacher's percent of steps completed and student's percent of maximum possible performance for social or academic behaviors.

**Student outcomes.** The students' responses were examined through visual analysis of the graphs generated from the data (Figure 10). All three students

demonstrated increased performance for their respective skills from baseline and IB to MA. Nonetheless, there was variability between students in performance consistency. The IB and MA means and standard deviations were taken for each student, and used to calculate Cohen's  $d$  (Table 2) to determine performance differences.

Agnes' average teacher points rating of her behavior was 7.7 ( $SD = 2.31$ ) at baseline, 5.3 ( $SD = 6.81$ ) at IB, and 13.7 ( $SD = 3.84$ ) at MA. During IB she demonstrated a downward trend. At MA she had an increased level in performance with variable ratings of her behavior in the classroom. She had a large improvement from IB to MA (Cohen's  $d = 1.51$ ). Agatha's average teacher points rating of her behavior was 20 ( $SD = 1$ ) at baseline, 25.23 ( $SD = 3.86$ ) at IB, and 26.5 ( $SD = 4.14$ ) at MA. During IB she demonstrated an upward trend and at MA she had an increased level in performance with variable ratings of her behavior in the classroom. She had a small improvement from IB to MA (Cohen's  $d = 0.30$ ). Jacqe's number of letters correctly identified was 24 at baseline. She had a mean of 32 ( $SD = 2.68$ ) letters named correctly at IB and a mean of 47 ( $SD = 2.68$ ) letters named correctly at MA. She had a large improvement from IB to MA (Cohen's  $d = 5.66$ ).

**Comparison of teacher and student data.** The comparisons of teacher and student data were conducted through visual analysis of the graphs generated from the data (Figure 11). All teachers demonstrated better performance at MA than IB, as all students demonstrated better performance at MA than IB. It appears that all the students had a better level of performance when their respective teachers implemented the intervention with increased accuracy.

Ms. Chief implemented the intervention at a consistent, but low level during IB, where Agnes showed a steady downward trend in performance. Ms. Chief showed a high percentage of accuracy during the PFDR, F1, F2, and F3 phases. Agnes's data indicated that she had a decreasing level of performance across these phases. This trend continued into the MA phase, despite Ms. Chief's continued high level of treatment integrity. Nonetheless, MS. Chief and Agnes had variable performance during the MA phase.

Ms. Sal implemented the intervention at a low level, but demonstrated an upward trend in treatment adherence during IB. Agatha showed an increase in her performance as Ms. Sal's treatment integrity increased in the IB phase. Ms. Sal demonstrated variability in the PFDR and F1 phases before consistently implementing the intervention. Agatha showed a decrease or greatest variability in her performance after low treatment integrity days. Across the F2 and F3 phases, Ms. Sal provided consistently accurate implementation of the intervention and Agatha showed an upward trend in points earned. Treatment integrity was at a high level throughout the MA phase and Agatha had her greatest level of points earned. Still, one data point was remarkably low for points earned when compared to the others, giving the student performance data variability.

Ms. Tiffye's treatment integrity was variable with a low level during the IB phase and Jacqe's performance for letters named correctly had a slight upward trend with minor variability. Ms. Tiffye showed perfect accuracy during the PFDR, F1, F2, and F3 phases and during these phases Jacqe had a steep upward trend in performance. Ms. Tiffye had a high level of treatment integrity with slight variability during the MA phase. Correspondingly, Jacqe continued to demonstrate an upward trend in performance with slight variability.

## CHAPTER V

### CONCLUSION

The purpose of this study was to demonstrate that treatment integrity can be maintained as the scheduled performance feedback procedures are reduced. Another goal was to effectively integrate indiscriminable contingencies into interventions for increasing teacher adherence to treatment protocols. Lastly, it was a secondary purpose to show treatment integrity and treatment integrity maintenance are associated with more desirable student outcomes.

#### **Treatment Integrity**

The primary purpose of this study was met. All three teachers demonstrated improved treatment integrity during maintenance than at implementation baseline. They were capable of performing the intervention as designed when receiving direct training; however, this did not carryover in to initial independent implementation, nor did providing performance feedback at the SAT meeting improve performance enough to meet the minimal standard of 80% intervention step adherence. All three teachers were engaged in the systematic performance feedback and direct rehearsal phase and went through the three fading phases. The teachers provided a high level of treatment integrity

during the maintenance phase (*range* = 88% to 99%) over an extended period of days (9 to 15 school days). Two of the three teachers presented minor variability in their adherence to the treatment protocol, and only one teacher demonstrated poor adherence to her intervention protocol during one day. Given the preponderance of data from all three teachers, the one day of poor performance may be considered aberrant.

### **Use of Indiscriminable Contingencies**

It appeared that the use of indiscriminable contingencies easily fit within in the teacher meetings. There were no ill effects mentioned by the teachers when specifically asked about the fading procedures. Ms. Chief remarked that the performance feedback helped her to focus on what was actually done and if it was working. Additionally, the use of indiscriminable contingencies helped to hold her accountable and make her aware of the intervention steps as well as motivate her to avoid negative feedback. Ms. Sal found it helpful to have structured support to help her with implementing a social behavior intervention. She found the graphic feedback to be helpful because it showed the correspondence between her treatment adherence and Agatha's earned points. She found the use of indiscriminable contingencies useful, in that she was relieved when she did not have to receive performance feedback and direct rehearsal because of missed steps. Ms. Tiffye stated that performance feedback was helpful to show her what steps she successfully and unsuccessfully completed. She too found that the indiscriminable contingencies helped to motivate stricter adherence to the intervention protocol. Overall, it appears that the teachers liked performance feedback because it allowed them to correct mistakes. Further, it appears that indiscriminable contingencies may have acted to cue negative reinforcement for accurately completing interventions by assuring greater

potential for less effort demanded from the teacher. Previous research supports the efficacy of negative reinforcement paradigms to increase teacher treatment integrity (Digennaro et al., 2005; Digennaro et al., 2007).

### **Impact on Student Outcomes**

Finally, it appears that treatment integrity has a positive impact on student performance, but this is not necessarily a one-to-one correspondence and the impact may vary based up the teacher and student dyad. All the students showed improved performance when their respective teacher had improved treatment integrity. Agnes and Jacqe had large level change in their performance when their respective teachers adhered to the intervention steps. Agatha showed a small improvement when her teacher increased her treatment integrity. It is possible that Agnes and Jacqe demonstrated greater deficits during the IB phase than Agatha and therefore had more room for improvement. Additionally, Ms. Sal may have completed the permanent products accurately, but the quality of step implementation might have been less than ideal. Agnes and Agatha demonstrated the greatest variability in their performance across the different phases of teacher training. They received interventions for social behavior, while Jacqe received an intervention for letter naming. Changes in social behavior, especially for children with extreme behaviors, may be more susceptible to a wider array of environmental cues than are addressed in any one intervention. This affirms the sentiment that treatment integrity is necessary, but insufficient to change student behavior (Gresham, 2005).

### **Contributions**

The primary investigator developed interventions for teachers to implement and measured treatment integrity through the completion of permanent products. The



permanent products were associated with behavioral expectations of teachers for each step of the intervention protocol, in a manner that could be observed, and measured over time. This is consistent with recommendations for measuring treatment integrity (Gresham, 1989; Gresham et al., 2000) and consulting in the schools (Sterling-Turner et al., 2002). The role of consultants within the school setting should include direct training and performance feedback to teachers regarding intervention implementation (Noell, 2008; Noell & Witt, 1996). The results appear to indicate that providing an intervention protocol and brief training is inadequate for satisfactory or consistent treatment integrity. Similarly, previous studies showed the greatest improvements for treatment integrity when teachers were given repeated practice with performance feedback (Coddling et al., 2005; Gilbertson et al., 2007; Noell et al., 1997; Sterling-Turner et al., 2001; Sterling-Turner et al., 2002).

The findings from this study and previous research support the use of direct rehearsal and performance feedback to aid teachers in acquiring skills to complete intervention steps; however, it is necessary for teachers to maintain these behaviors over time. Some studies showed promise for practice and feedback to increase treatment integrity maintenance (Carnine & Fink, 1978; Coddling et al., 2005), but further investigations found the maintenance of teacher treatment integrity to be variable (Gilbertson et al., 2007; Noell et al., 1997; Noell et al., 2000; Witt et al., 1997) or diminish over time (Duhonet et al., 2009; Reinke et al., 2008). Attempts to thin feedback sessions have shown promise, but resulted in variable treatment integrity and call for more systematic inquiry (Noell et al., 2000; Noell et al., 2002).

Similar to a study by Digennaro et al. (2007), this study sought to employ systematic fading from daily performance feedback to once every other week. This was proposed to work by employing negative reinforcement through avoidance of meetings. Unfortunately, these researchers were unable to fade to every other week for three of four participants, but the one participant who completed the fading process demonstrated high integrity and low variability (Digennaro et al., 2007). This study incorporated meeting reduction and fading, but omitted multiple pre-fading phases. An additional component of indiscriminable contingency was added to the fading phases because it would reduce the participants' ability to discriminate the treatment phase from the maintenance phase, which should further contribute to behavior maintenance (Freeland & Noell, 2002). Most importantly, the teachers in this study were able to complete the fading procedures and then were monitored for at least two school weeks. The resulting data indicated that all three teachers were able to adhere to the treatment steps over time with a high level of accuracy and slight variability.

### **Limitations**

Despite the promising results from this study it is not without limitations. The first limitation is the use of a single-case design, which makes it difficult to generalize the results beyond teachers and students similar to the participants. Future research should employ group designs to determine the impact of this fading process across samples with more diverse demographics. Another short coming of this study was that one teacher did not receive a SAT meeting during the IB phase because of the student's volatile and aggressive behaviors. While it would have been ethically remiss to allow the student to continue with these actions, it limits the confidence in the conclusion that teacher

behavior has limited change with feedback given in a SAT meeting. The literature may benefit from continued single-case and group design research that examines the impact of feedback given during the SAT process.

The reliance on permanent products may provide challenges such as discerning the quality of treatment adherence, or if teachers were reinforced for product completion or treatment adherence. Researchers should allocate their efforts to measure the accuracy of permanent products as related to treatment integrity. Also, relying on teacher rating of student behavior may provide difficulties as unknown environmental factors might influence the perceptions of the teacher. Researchers should take precautions to guard against these influences and their impact on the quality of treatment integrity. Further investigations are needed to determine if treatment integrity impacts social and academic behaviors differently. Lastly, researchers should investigate how the behaviors of the consultant increase or decrease teacher receptiveness to participate in activities such as performance feedback with direct rehearsal.

In conclusion, the data from this study indicates that there is potential for the use of systematic fading with indiscriminable contingencies to improve teacher treatment integrity. These improvements appear to be sustained over at least two weeks time and teacher adherence to intervention protocols seems stable. Improvements in treatment adherence may have positive effects for students, but this may vary depending upon teacher, student, and teacher-student dyad needs and characteristics. Researchers should continue to explore the arena of treatment integrity to determine how to best improve it and how it impacts student behaviors.

## REFERENCES

- Baer, D., Wolf, M., & Risely, R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97.
- Bahr, M. & Kovalski, J. (2006). The need for problem-solving teams. *Remedial and Special Education, 27*, 2-5.
- Borelli, B., Sepinwall, D., Ernst, D., Bellg, A., Czajkowski, S., Breger, R., DeFrancesco, C, Levesque, C, Sharp, D., Ogedegbe, G., Resnick, B., & Orwig, D. (2005). A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of Consulting and Clinical Psychology, 73*, 852-860.
- Burke, M. & Hagan-Burke, S. (2007). Concurrent criterion-related validity of early literacy indicators for middle of first grade. *Assessment for Effective Intervention, 32*, 66-77.
- Burns, M., Vanderwood, M., & Ruby, S. (2005). Evaluating the readiness of pre-referral intervention teams for use in a problem solving model. *School Psychology Quarterly, 20*, 89-105.

- Carnine, D. & Fink, W. (1978). Increasing the rate of presentation and use of signals in elementary classroom teachers. *Journal of Applied Behavioral Analysis, 11*, 35-46.
- Carr, J. (2005). Recommendations for reporting multiple-baseline designs across participants. *Behavioral Interventions, 20*, 219-224.
- Codding, R., Feinberg, A., Dunn, E., & Pace, G. (2005). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavioral Analysis, 38*, 205-219.
- Cooper, J., Heron, T., & Heward, W. (2007). *Applied Behavior Analysis* (2<sup>nd</sup> ed.). Upper Saddle River, New Jersey: Pearson Education, Inc.
- DiGennaro, F., Martens, B., & Kleinmann, A. (2007). A comparison of performance feedback procedures on teachers' treatment implementation integrity and students inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis, 40*, 447-461.
- DiGennaro, F., Martens, B., & McIntyre, L. (2005). Increasing treatment integrity through negative reinforcement: Effects on teacher and student behavior. *School Psychology Review, 34*, 22-31.
- Duhon, G., Mesmer, E., Gregerson, L., & Witt, J. (2009). Effects of public feedback during RTI team meetings on teacher implementation integrity and student academic performance. *Journal of School Psychology, 47*, 19-37.
- Duhon, G.J., Noell, G.H., Witt, J.C., Freeland, J.T., Dufrene, B.A. & Gilbertson, D.N. (2004). Identifying academic skill and performance deficits: The experimental

- analysis of brief assessments of academic skills. *School Psychology Review*, 33, 429-443.
- Federal Interagency Forum on Child and Family Statistics. (2009). *America's Children: Key National Indicators of Well-Being, 2009*. Washington, DC: U.S. Government Printing Office.
- Freeland, J. & Noell, G. (2002). Programming for maintenance: An investigation of delayed intermittent reinforcement and common stimuli to create indiscriminable contingencies. *Journal of Behavioral Education*, 11, 5-18.
- Gilbertson, D., Witt, J., LaFleur-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.
- Gresham, F. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review*, 18, 37-50.
- Gresham, F. (2005). Treatment integrity and therapeutic change: Commentary on Perepletchikova and Kazdin. *Clinical Psychology: Science and Practice*, 12, 391-394.
- Gresham, F., MacMillian, D., Beebe-Frankeberger, M., & Bocian, K. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research and Practice*, 15, 198-205.
- Griffith, A., Duppong-Hurley, K., & Hagaman, J. (2009). Treatment integrity of literacy

- interventions for students with emotional and/or behavior disorders. *Remedial and Special Education, 30*, 245-255.
- Guevremont, D., Osnes, P., & Stokes, T. (1986). Programming maintenance after correspondence training interventions with children. *Journal of Applied Behavioral Analysis, 19*, 215-219.
- Hagermoser-Sanetti, L. & Kratochwill, T. (2009). Treatment integrity assessment in the schools: An evaluation of the treatment integrity planning protocol. *School Psychology Quarterly, 24*, 24-35.
- McDougal, J., Nastasi, B., & Chafouleas, S. (2005). Bringing research into practice to intervene with young behaviorally challenging students in public school settings: Evaluation of the behavioral consultation team (BCT) project. *Psychology in the Schools, 42*, 537-551.
- McNamara, K., Rashee, H., & DeLamatre, J. (2008). A statewide study of school-based intervention teams: Characteristics, member perceptions, and outcomes. *Journal of Educational and Psychological Consultation, 18*, 5-30.
- Moore, J., Edwards, R., Sterling-Turner, H., Riley, J., DuBard, M., & McGeorge, A. (2002). Teacher acquisition of functional analysis methodology. *Journal of Applied Behavioral Analysis, 35*, 73-77.
- Moncher, F. & Prinz, R. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11*, 247-266.
- Noell, G. (2008). Appraising and praising systematic work to support systems change: Where we might be and where we might go. *School Psychology Review, 37*, 333-336.

- 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. & Gresham, F. (2001). A multiple-sequence variant of the multiple-baseline design: A strategy for analysis of sequence effects and treatment comparison. *School Psychology Quarterly, 16*, 207-221.
- Noell, G. & Witt, J. (1996). A critical re-evaluation of five fundamental assumptions underlying behavioral consultation. *School Psychology Quarterly, 11*, 189-203.
- Noell, G. & Witt, J. (1999). When does consultation lead to intervention? Critical issues for research and practice. *The Journal of Special Education, 33*, 29-35.
- Noell, G., Witt, J., Gilbertson, D., Ranier, D., & Freedland, 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 Behavioral Analysis, 33*, 271-284.
- Noell, G., Witt, J., Slider, N., Connell, J., Gatti, S., Williams, K...Duhon, G. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34*, 87-106.
- Perepletchikova, F. & Kazdin, A. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12*, 365-394.



- Reinke, W., Lewis-Palmer, T., & Martin, E. (2007). The effects of visual performance feedback on teacher use of behavior-specific praise. *Behavior Modification, 31*, 247-263.
- 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.
- Shaywitz, S., Morris, R. & Shaywitz, B. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology, 59*, 451-475.
- Slonski-Fowler, K. & Truscott, S. (2004). General education teachers' perceptions of the prereferral intervention team process. *Journal of Educational and Psychological Consultation, 15*, 1-39.
- Smith, S., Daunic, A., & Taylor, 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.
- Snyder, T., Dillow, S., & Hoffman, C. (2009). *Digest of Education Statistics 2008 (NCES 2009-020)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Stage, S., Jackson, H., Jensen, M., Moscovitz, K., Bush, J., Violette, H., Thurman, S., Olson, E., Bain, N., & Pious, C. (2008). A validity study of functionally-based behavioral consultation with students with emotional/behavioral disabilities. *School Psychology Quarterly 23*, 327-353.
- Sterling-Turner, H., Watson, T., & Moore, J. (2002). The effects of direct training and

- treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly*, 17, 47-77.
- Sterling-Turner, H., Watson, T., Wildmon, M., Watkins, C. & Little, E. (2001). Investigating the relationship between training type and treatment integrity. *School Psychology Quarterly*, 16, 56-67.
- Stokes, T. & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavioral Analysis*, 10, 349-367.
- Tawney, J. & Gast, D. (1984). *Single Subject Research in Special Education*. Columbus, OH: Merrill.
- Truscott, S., Cohen, C., Sams, D., Sanborn, K. & Frank, A. (2005). The current state(s) of prereferral intervention teams. *Remedial and Special Education*, 26, 130-140.
- Ward, P., Johnson, M., & Konukman, F. (1998). Direct rehearsal and preservice teachers' performance of instructional behaviors. *Journal of Behavioral Education*, 8, 369-380.
- Witt, J., Noell, G., LeFleur, L., & Mortenson, B. (1997). Teacher use of interventions in general education settings: Measurement and analysis of independent variable. *Journal of Applied Behavioral Analysis*, 30, 693-696.
- Wickstrom, K., Jones, K., LaFleur L., & Witt, J. (1998). An Analysis of Treatment Integrity in School-Based Behavioral Consultation. *School Psychology Quarterly*, 13,141-154
- Wood, B., Umbreit, J., Liaupsin, C., & Gresham, F. (2007). A treatment integrity analysis of function-based intervention. *Education and Treatment of Children*, 30, 105-120.

Zisimopoulos, D. (2010). Enhancing Multiplication Performance in Students with Moderate Intellectual Disabilities Using Pegword Mnemonics Paired with a Picture Fading Technique. *Journal of Behavioral Education, 19*, 117-133.

## APPENDICES

Figure 1

Agnes Check-in

Materials Needed:

Daily Chart

Pencil/Pen

Stickers and sticker chart

Procedure:

1. Ms. Chief will write the date in the feedback chart.

Before each class period (8:30) Agnes will check-in with Ms. Chief. She will let her know what her expectations are and check each expectation as they are told to her.

2. Sitting

3. Listening

4. Participating

After the each segment (every 10 minutes): Ms. Chief will meet with Agnes and circle the corresponding number/face for how she did and provide a brief explanation why she received the points she did.

5. Sitting: 0-2 for each 10 minute period is circled

6. Listening: 0-2 for each 10 minute period is circled

7. Participating: 0-2 for each 10 minute period is circled

8. Sitting: Score is totaled and explained to her (6-5 = Great, 4 = OK, 3-0 = Poor)

9. Listening: Score is totaled and explained to her (6-5 = Great, 4 = OK, 3-0 = Poor)

10. Participating: Score is totaled and explained to her (6-5 = Great, 4 = OK, 3-0 = Poor)

11. Ms. Chief will write down the morning meeting total points and briefly review with her what is meant by the points she earned. Check the appropriate point line; let her know the number of stickers she earned.

12. If she earned stickers have her place it on her sticker chart.

\*\*After 8 stickers, she earns a prize

Expectations	Sitting	Listening	Working
Definition	Bottom in her chair	Eyes and ears on the teacher; mouth quiet	Assignments completed

2 points = 0-1 redirection

1 point = 2-3 redirections

0 points = 4+ redirections

Figure 2

Date:	Sitting	Listening	Working
8:30 – 8:40	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
Beginning	0 1 2	0 1 2	0 1 2
8:40 – 8:50	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
Middle	0 1 2	0 1 2	0 1 2
8:50 – 9:00	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
End	0 1 2	0 1 2	0 1 2
My Total =	Sitting =	Listening =	Working =

Point Total:

\_\_\_\_\_ 18 - 14 = ☺ day (2 stickers)  
 \_\_\_\_\_ 13 - 10 = so-so day (1 sticker)  
 \_\_\_\_\_ 9 - 0 = ☹ day (0 stickers)

Points Expectations:

6-5 = Great  
 4 = OK  
 3-0 = Poor

Date:	Sitting	Listening	Working
8:30 – 8:40	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
Beginning	0 1 2	0 1 2	0 1 2
8:40 – 8:50	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
Middle	0 1 2	0 1 2	0 1 2
8:50 – 9:00	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
End	0 1 2	0 1 2	0 1 2
My Total =	Sitting =	Listening =	Working =

Point Total:

\_\_\_\_\_ 18 - 14 = ☺ day (2 stickers)  
 \_\_\_\_\_ 13 - 10 = so-so day (1 sticker)  
 \_\_\_\_\_ 9 - 0 = ☹ day (0 stickers)

Points Expectations:

6-5 = Great  
 4 = OK  
 3-0 = Poor

Date:	Sitting	Listening	Working
8:30 – 8:40	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
Beginning	0 1 2	0 1 2	0 1 2
8:40 – 8:50	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
Middle	0 1 2	0 1 2	0 1 2
8:50 – 9:00	☹ 1 ☺	☹ 1 ☺	☹ 1 ☺
End	0 1 2	0 1 2	0 1 2
My Total =	Sitting =	Listening =	Working =









Point Total:









\_\_\_\_\_ 18 - 14 = ☺ day (2 stickers)  
 \_\_\_\_\_ 13 - 10 = so-so day (1 sticker)  
 \_\_\_\_\_ 9 - 0 = ☹ day (0 stickers)

Points Expectations:

6-5 = Great  
 4 = OK  
 3-0 = Poor

Figure 3

1	2	3	4
			
5	6	7	8
			

1	2	3	4
			
5	6	7	8
			









1	2	3	4
			
5	6	7	8
			

Figure 4

Agatha Check-in

Materials Needed:

- Daily Chart
- Pencil/Pen
- Stickers and sticker chart

Procedure:

1. Ms. Sal will write the date in the feedback chart and make sure
2. Agatha's area is clearly marked on the floor around her desk.

Before morning work, Agatha will check-in with Ms. Sal: She will let Agatha know what her expectations are and check each expectation as they are told to her for:

3. Stay in Your Area
4. Work Quietly
5. Focus on Your Work

After the Math period: Ms. Werchan will meet with Desiree and provide a brief explanation why she received the points she did.

6. Stay in Your Area: 0-2 for each class period is circled immediately after that period
7. Work Quietly: 0-2 for each class period is circled immediately after that period
8. Focus on Your Work: 0-2 for each class period is circled immediately after that period
9. Stay in Your Area: Score is totaled and explained to her (10-9 = Great, 8-7 = OK, 6-0 = Poor)
10. Work Quietly: Score is totaled and explained to her (10-9 = Great, 8-7 = OK, 6-0 = Poor)
11. Focus on Your Work: Score is totaled and explained to her (10-9 = Great, 8-7 = OK, 6-0 = Poor)
12. Ms. Sal will write down the morning total points and briefly review with her what is meant by the points she earned. Check the appropriate point line; let her know the number of stickers she earned.
13. If she earned stickers have her place it on her sticker chart.

\*\*After 8 stickers, she earns 10 extra dollars

Expectations Chart

Expectations Period	Stay in Your Area: her desk and body are within the taped area on the floor or designated area in the hallway	Work Quietly: works without engaging peers in conversations, she is silent to show focus on assignments	Focus on Your Work: she has all items of her assignments completed and class participation is appropriate
Morning Work	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 90%-100% complete 1 = 89% - 70% complete 0 = less than 70%
Bathroom Break	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 90%-100% complete 1 = 89% - 70% complete 0 = less than 70%
Spelling	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 90%-100% complete 1 = 89% - 70% complete 0 = less than 70%
Language Arts	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 90%-100% complete 1 = 89% - 70% complete 0 = less than 70%
Math	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 2 or less redirections 1 = 3 - 4 redirections 0 = 4 or more	2 = 90%-100% complete 1 = 89% - 70% complete 0 = less than 70%



Figure 5

Date:	Stay in Your Area	Work Quietly	Focus on Your Work
Morning Work	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Bathroom Break	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Spelling	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Language Arts	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Math	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
A.M. Total			

Point Total:

\_\_\_\_\_ 30 - 27 = ☺ day (2 stickers)  
 \_\_\_\_\_ 26 - 21 = so-so day (1 sticker)  
 \_\_\_\_\_ 20 - 0 = ☹ day (0 stickers)

Points Expectations:

10-9 = Great  
 8-7 = OK  
 6-0 = Poor

Date:	Stay in Your Area	Work Quietly	Focus on Your Work
Morning Work	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Bathroom Break	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Spelling	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Language Arts	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
Math	☹ 0 1 2 ☺	☹ 0 1 2 ☺	☹ 0 1 2 ☺
A.M. Total			

Point Total:

\_\_\_\_\_ 30 - 27 = ☺ day (2 stickers)  
 \_\_\_\_\_ 26 - 21 = so-so day (1 sticker)  
 \_\_\_\_\_ 20 - 0 = ☹ day (0 stickers)

Points Expectations:

10-9 = Great  
 8-7 = OK  
 6-0 = Poor

Figure 6

### Jacqe Letter Naming with Error Correction and Reward

#### Materials:

Sheet with alphabet boxed-lists

Pencil

Sticker Chart

#### Cold Run:

1. Present the letters sheet to Jacqe and randomly point to each letter, one at a time. She has 3 seconds to name the upper case letter you are pointing at.
2. If she names it within 3 seconds, make a slash-mark in the box under the "c" heading. Then point to the next letter. If she gives the wrong letter name, takes longer than 3 seconds, or says "I don't know," tell her the name of the letter, and have her repeat it. Make circle in the box under the "c" heading.
3. Write the number of letters she named correctly under the letter list "c" heading.
4. Jacqe will receive a sticker for naming more letters correct on the cold run than the previous cold run.

#### Practice with feedback:

5. With the incorrect letters, point to them one at a time in alphabetical order and say the letter name and have her repeat the letter name while looking at the letters. Mark upper left corner of the letter box.
6. With the incorrect letters, point to them one at a time in reverse alphabetical order and say the letter name and have her repeat the letter name while looking at the letters. Mark upper right corner of the letter box.
7. Randomly point to each letter one at a time and say the letter name and have her repeat the letter name while looking at the letters. Mark lower left corner of the letter box.
8. Randomly point to each letter one at a time and say the letter name and have her repeat the letter name while looking at the letters. Mark lower right corner of the letter box.

#### Hot Run:

9. Present the letters sheet to Jacqe and point to each letter, one at a time. She has 3 seconds to name the upper case letter you are pointing at.
10. If she names it within 3 seconds, make a slash-mark in the box under the "h" heading. Then point to the next letter. If she gives the wrong letter name, takes longer than five seconds, or says "I don't know," tell her the name of the letter, and have her repeat it. Make circle in the box under the "c" heading.
11. Write the number of letters she named correctly under the letter list "h" heading.
12. Jacqe will receive a sticker for naming more letters correct on the hot run than the session's cold run.

\*\*After 8 stickers, Jacqe get to pick a prize.

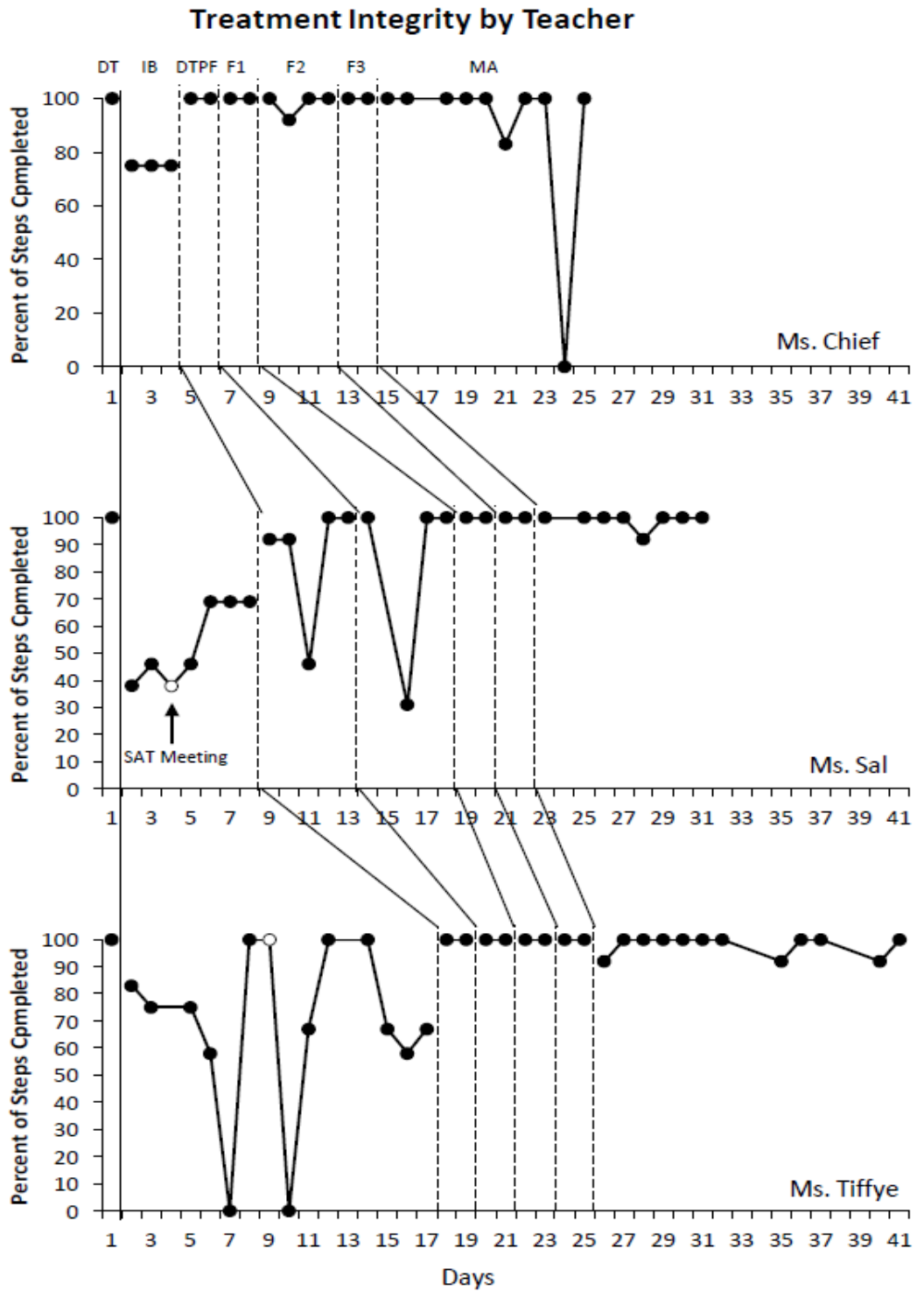
Figure 7

c	h		c	h		c	h		c	h		c	h	
		A			A			A			A			A
		B			B			B			B			B
		C			C			C			C			C
		D			D			D			D			D
		E			E			E			E			E
		F			F			F			F			F
		G			G			G			G			G
		H			H			H			H			H
		I			I			I			I			I
		J			J			J			J			J
		K			K			K			K			K
		L			L			L			L			L
		M			M			M			M			M
		N			N			N			N			N
		O			O			O			O			O
		P			P			P			P			P
		Q			Q			Q			Q			Q
		R			R			R			R			R
		S			S			S			S			S
		T			T			T			T			T
		U			U			U			U			U
		V			V			V			V			V
		W			W			W			W			W
		X			X			X			X			X
		Y			Y			Y			Y			Y
		Z			Z			Z			Z			Z

Figure 8

c	h		c	h		c	h		c	h		c	h	
		a			a			a			a			a
		b			b			b			b			b
		c			c			c			c			c
		d			d			d			d			d
		e			e			e			e			e
		f			f			f			f			f
		g			g			g			g			g
		h			h			h			h			h
		i			i			i			i			i
		j			j			j			j			j
		k			k			k			k			k
		l			l			l			l			l
		m			m			m			m			m
		n			n			n			n			n
		o			o			o			o			o
		p			p			p			p			p
		q			q			q			q			q
		r			r			r			r			r
		s			s			s			s			s
		t			t			t			t			t
		u			u			u			u			u
		v			v			v			v			v
		w			w			w			w			w
		x			x			x			x			x
		y			y			y			y			y
		z			z			z			z			z

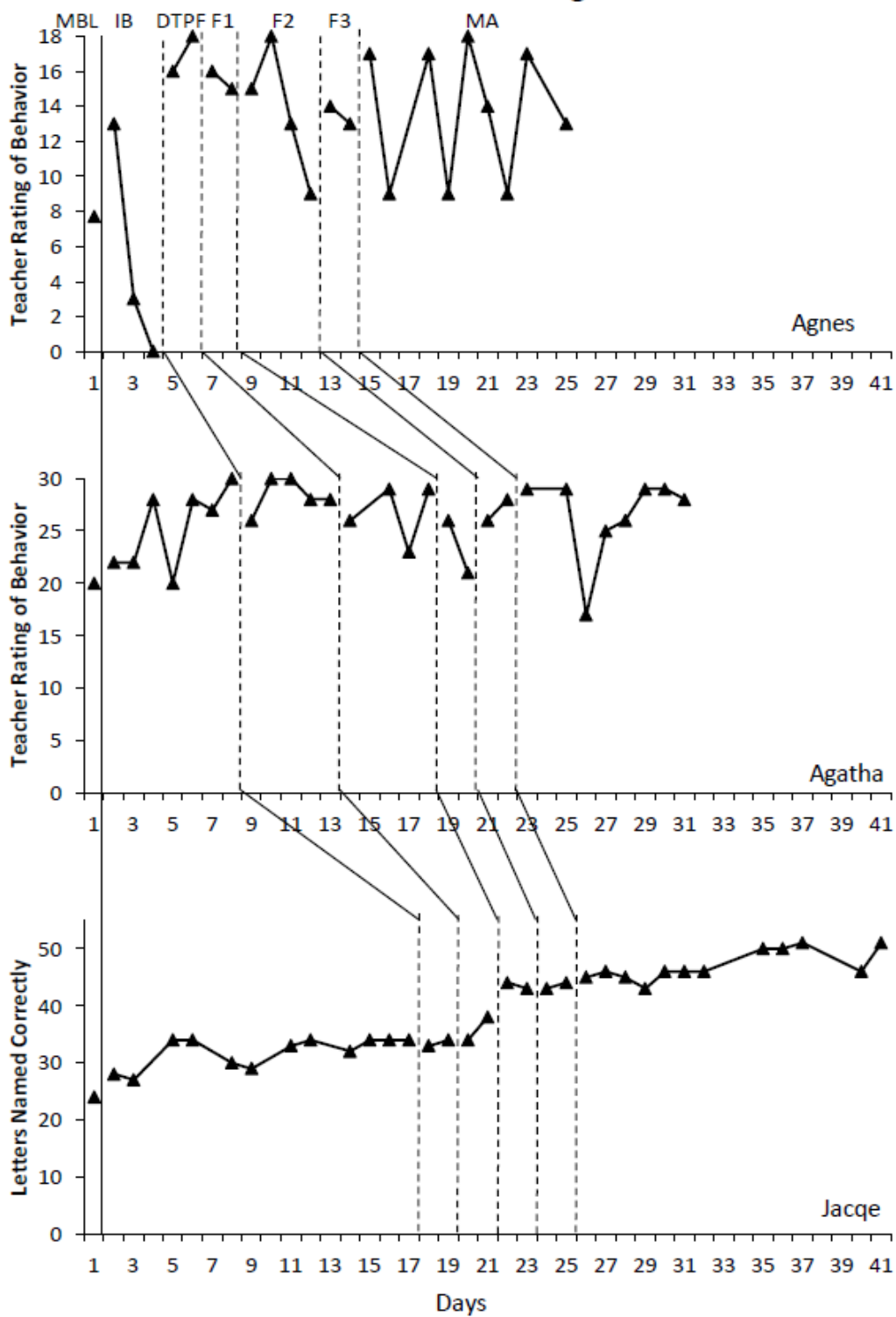
Figure 9



DT = Direct Training; IB = Implementation Baseline; PFDR = Performance Feedback Direct Rehearsal; F1 = Fading 1; F2 = Fading 2; F3 = Fading 3; MA = Maintenance

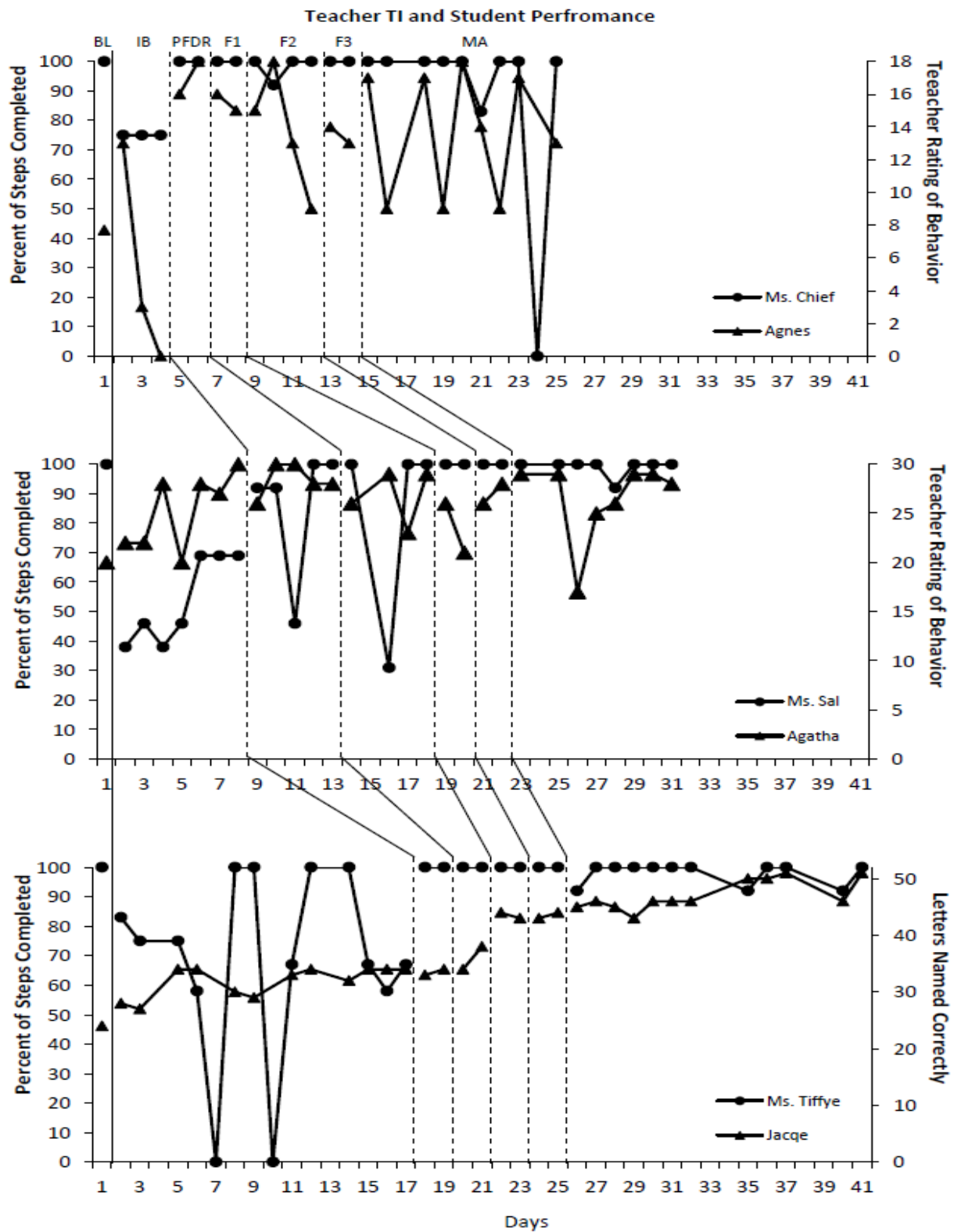
Figure 10

Student Progress



MBL= Mean Baseline Performance; IB = Implementation Baseline; PFDR = Performance Feedback Direct Rehearsal; F1 = Fading 1; F2 = Fading 2; F3 = Fading 3; MA = Maintenance

Figure 11



BL = Students' Mean Baseline Performance/Teachers' Direct Training Performance; IB = Implementation Baseline; PFDR = Performance Feedback Direct Rehearsal; F1 = Fade 1; F2 = Fade 2; F3 = Fade 3; MA = Maintenance

Table 1. Mean Percentage of Correct Steps

Teacher	IB <sup>a</sup>		MA <sup>b</sup>		Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Ms. Chief	75	0	88	31.48	0.60
Ms. Sal	54	14.80	99	2.83	4.26
Ms. Tiffye	68	32.66	98	3.62	1.3

<sup>a</sup>Mean of all implementation baseline points; <sup>b</sup>Mean of all maintenance points.



Table 2. Mean Student Performance

Teacher	IB <sup>a</sup>		MA <sup>b</sup>		Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Agnes <sup>c</sup>	5.33	6.81	13.67	3.84	1.51
Agatha <sup>c</sup>	25.23	3.86	26.5	4.14	0.30
Jacqe <sup>d</sup>	31.92	2.68	47.08	2.68	5.66

<sup>a</sup>Mean of all baseline points; <sup>b</sup>Mean of all maintenance points.

<sup>c</sup>Teacher rating in total points; <sup>d</sup>Letters named correctly

## VITA

Thomas Jai Gross

Candidate for the Degree of

Doctor of Philosophy

Thesis: ENHANCING TREATMENT INTEGRITY THROUGH SYSTEMATIC  
FADING WITH INDISCRIMINABLE CONTINGENCIES

Major Field: Educational Psychology, Specialization in School Psychology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Educational Psychology, Specialization in School Psychology at Oklahoma State University, Stillwater, Oklahoma by December, 2013.

Completed the requirements for the Master of Arts in Community Counseling at Marquette University, Milwaukee, WI in 2004.

Completed the requirements for the Bachelor of Arts in Psychology at Marquette University, Milwaukee, WI in 2002.

Experience:

Heart of America Psychology Training Consortium/Royal Oaks Hospital. Windsor, MO. Pre-Doctoral Psychology Intern.

Milwaukee Center for Independence/Birth to Three. Milwaukee, WI  
Psychotherapist.

Aurora Psychiatric Hospital. Wauwatosa, WI. Group leader/Therapist

Professional Memberships:

American Psychological Association – Student Member

National Association of School Psychologists – Student Member