2013

Treatment integrity failures matched to behavioral function

Joanna Elizabeth Lomas Mevers
Louisiana State University and Agricultural and Mechanical College, joanna.lomasmevers@choa.org

TREATMENT INTEGRITY FAILURES MATCHED TO BEHAVIORAL FUNCTION

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Joanna E. Lomas Mevers
B.S., Louisiana State University, 2001
M.A., Georgia State University, 2008
August 2013

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Prior research on treatment integrity has focused on either the lack of measurement of the independent variable (Peterson, Homer & Wonderlich, 1982; Gresham, Gansle & Noel, 1993; Wheeler, Baggett, Fox & Blevins, 2006; McIntyre, Gresham, DiGennaro & Reed, 2007; Sanetti, Gritter & Dobey, 2011) or on methods to increase overall levels of treatment integrity(Witt, Noell, LaFleur & Mortenson, 1997; Noell, Witt, Gilbertson, Ranier & Freeland,1997; Noell et al., 2005). Yet little research has been devoted to understanding the effectiveness of common interventions when those interventions are implemented with less than perfect integrity. The current investigation evaluated the effectiveness of using reinforcement and prompting to increase correct item completion on math worksheets for kindergarten and first graders. Treatment was evaluated when both components were implemented, when only reinforcement was implemented, when only prompting was implemented and when neither was implemented. In addition preferences for either attention or escape were evaluated as moderator variables to understand how individual differences impact treatment effectiveness. Results indicated treatment was effective at all levels of implementation when moderator variables were not accounted for. However when moderator variables were evaluated individuals who preferred escape responded best when both treatment components were implemented whereas, for individuals who preferred attention all treatment conditions were equally effective.

INTRODUCTION

To date, there is an abundance of behavioral research that has focused on design and selection of effective interventions (Brosnan & Healy, 2011; Gresham, 2011; Iwata, Pace, Cowdery, & Miltenberger, 1994; Wood, Blair, & Ferro, 2009). Intervention selection has commonly focused on the use of functional behavior assessments (FBA) and functional analysis (FA) as tools to gather information to select effective interventions that are likely to target the maintaining reinforcer. Function based interventions are designed to disrupt the relationship between the identified functional reinforcer and the target behavior through the use of extinction and to teach a more appropriate replacement behavior. The alternative replacement behavior allows the individual to continue to access the functional reinforcer, but in a more socially appropriate manner (Carr & Durand, 1985). This combination of extinction and reinforcement often leads to both a decrease in problem behavior as well as an increase in appropriate behavior. Intervention selection has also relied on the use of empirically validated interventions that have proven to be effective to address specific problems. For example, The National Reading Panel (2000) published a report outlining reading interventions that have proven to be effective at targeting the underlying deficits that lead to poor reading achievement. Both function based interventions and empirically validated interventions are established methods for identifying effective treatments that are likely to lead to positive outcomes for children.

There is no shortage of research focused on identifying interventions to address a variety of specific problem behaviors, such as off-task behavior (Austin & Soedo, 2008; Flood, Wilder, Flood & Masuda, 2002; Stahr, Cushing, & Lane, 2006), skill acquisition (Freeland &

Noell, 1999; Mayfield & Vollmer, 2007; Noell, Connell, & Duhon, 2006), aggression (Lalli & Casey, 1996; Thompson, Fisher, Pizza, & Kuhn,1998; Worsdell, Iwata, Hanely, Thompson, & Kahng, 2000), self-injurious behavior (Fischer, Iwata, & Mazaleski, 1997; Poling & Normand, 1999; Vollmer, Marcus, & Ringdahl, 1995), and disruptive behaviors (Greene, Bailey, & Barber, 1981; Jones, Drew, & Weber, 2000; Schieltz et al., 2011). Behavioral interventions have also been developed in a variety of contexts such as the home (Derby, et al., 1997; O'Brein, Riner, & Budd,1983; Wahler, Vigilante, & Strand, 2004), school (Campbell & Anderson, 2011; Kleinman & Saigh, 2011; Sasso et al., 1992), clinic (Kahng, Iwata, DeLeon, & Worsdell, 1997; Piazza et al., 1999; Roane, Fisher, & Sgro, 2001), and hospital (Cataldo, Bessman, Parke, Pearson, & Rogers, 1979; Fisher et al., 1993; Ingham & Andrews, 1973), and with a variety of populations (e.g. developmental disabilities, Attention Deficit Hyperactivity Disorder [ADHD], typically-developing, autism) (DuPaul, Ervin, Hook, & McGoey, 1998; Gardner, Wacker, & Boelter, 2009; Hagopian, Kuhn, & Strother, 2009; Kennedy, Meyer, Knowles, & Shukla, 2000). Although function based behavioral treatment research has examined varied concerns, contexts, and populations, most studies evaluate effectiveness under perfect or near perfect conditions, because the nature of research is to control as many variables as possible. These conditions often include the use of analog settings, with interventionists that are highly trained and implement the interventions with high treatment integrity. Although these conditions are ideal for understanding how effective a specific intervention is under research conditions, they lack the ability to evaluate how effective the same intervention is when implemented under less than ideal circumstances.

Some research has sought to understand the effects of implementation under less ideal circumstances, however these studies are limited in the behavior analytic literature and are therefore in need of more attention. This limited research threatens the generalizability of the intervention's effectiveness. Generalization refers to the extent to which the results of a study will be the same across different settings, individuals, and contexts and indicates the extent to which interventions will work as well in homes and schools, where parents and teachers are the interventionists. Unlike an analog environment where the interventionist only has to implement the intervention, in a home or school, parents and teachers are presented with competing demands (e.g., other children to tend to, dinner to cook), competing reinforcers (e.g. watching tv, talking with other teachers) and other people that may not be aware or trained to implement the intervention. To date the literature has addressed some of these threats by conducing replications in more natural environments and including the use of parents and teachers as interventionists (Campbell & Anderson, 2011; Derby et al., 1997; Kleinman & Saigh, 2011; O'Brien, Riner, & Budd,1983; Sasso et al., 1992; Wahler et al., 2004). These lines of research have provided some knowledge on individual intervention durability; however, what we know in this area is still very limited.

One area that has received little attention is the role treatment integrity plays in treatment generalizability and treatment effectiveness. Treatment integrity is the degree to which a treatment is implemented as intended (Peterson et al., 1982; Yeaton & Sechrest, 1981; Gresham, 1989; Salend, 1983; Watson, Sterling & McDade, 1997). Treatment integrity is important from both an internal validity perspective and an external validity perspective, and thus should be the focus of more research. Internal validly refers to the extent to which

changes in the dependent variable are due to manipulation of the independent variable and not due to extraneous variables. Therefore it is important to measure, define, monitor, and manipulate the independent variable (treatment implementation) so that researchers can draw valid conclusions regarding their data (Johnston & Pennypacker, 1980).

To date, researchers have devoted some energy to addressing different facets of treatment integrity which fall into three general categories. These three categories include research focused on 1) increasing reporting and measuring of independent variables, 2) increasing the likelihood of high integrity implementation, and 3) intervention effectiveness when treatment integrity is low. All of these research lines need more attention, and each represents a different way in which treatment integrity affects the validity of behavioral interventions.

Treatment Integrity in Experimental Research

Poor and unknown treatment integrity threatens the very nature of research (Gresham, Gansle, & Noell, 1993; Peterson et al., 1982; Yeaton & Sechrest, 1981), yet reporting of treatment integrity data is rare in the behavior analytic and school psychology literature (Peterson et al.; Gresham, Gansle & Noell, 1993; Wheeler et al., 2006; McIntyre, Gresham, DiGennaro, & Reed, 2007; Sanetti, et al., 2011). The problem may be due to the assumption that there is high treatment integrity in research studies, but this assumption does not align with behavior analytic epistemology. Behavior analysis relies on the demonstration of the functional relationship between the independent and dependent variables (Baer, Wolf, & Risley, 1968) and is necessary for replications of published studies. Demonstration requires defining, measuring, and monitoring treatment implementation.

The lack of reported treatment integrity data has been repeatedly documented in the literature and was first reported by Peterson et al. (1982). They reviewed 539 studies published in *The Journal of Applied Behavior Analysis (JABA)* from 1968 to 1980. They divided articles into three categories, (1) those that reported treatment integrity, (2) those that did not report treatment integrity data but had low risk for treatment integrity errors (e.g., machine delivered reinforcement, permanent products) or (3) those that did not report treatment integrity data and were considered at risk for treatment integrity errors. The results indicated that on average only 20% of the included studies reported measurement of the independent variable, and the authors coined the term "a curious double standard" to reflect this inconsistency with the behavior analytic approach. That is, there is a heavy emphasis in behavior analytic research on measurement of the dependent variable with little emphasis on measurement of the independent variable.

In a second study by Gresham, Gansle and Noell (1993), the authors reviewed 158 studies published between 1980 and 1990. The included studies came from a variety of behaviorally-oriented journals including; *Behavior Modification, Behavior Therapy, Journal of Applied Behavior Analysis (JABA), Journal of Abnormal Child Psychology, Journal of Consulting and Clinical Psychology, Journal of Behavior Therapy and Experimental Psychiatry and Behavioral Disorders.* The results of their study replicated the findings of Peterson et al. (1982). They found that only 16% of studies included studies reported data on treatment integrity, indicating that there were no improvements in rates of reporting data on implementation of the independent variable in the ten years since the publication of the Peterson et al. study.

This pattern of low levels of reported treatment integrity has remained low with only small improvements and was again demonstrated by Wheeler et al. (2006) and McIntyre et al. (2007). Wheeler et al. reviewed 60 studies on autism published from 1993 to 2003. The majority of included articles came from *JABA* (n=36) and the remaining articles came from *Focus on Autism and Other Developmental Disabilities, Journal of Positive Behavior Interventions, Research in Developmental Disabilities, Journal of Autism and Developmental Disorders, Education and Treatment of Children, Education and Training in Mental Retardation and Developmental Disabilities, Journal of Early Intervention, and Journal of Developmental and Physical Disabilities*. Results found that only 18% of included studies reported treatment integrity data. In the McIntye, et al. study they reviewed 152 studies published in *JABA* from 1991 to 2005 and found that only 30% of the included studies reported treatment integrity data.

More recently Sanetti, et al. (2011) published a study reviewing 72 studies in the school psychology literature from 1995 to 2008. Articles came from *Journal of School Psychology, Psychology in Schools, School Psychology Quarterly,* and *School Psychology Review.* The majority of the studies reviewed targeted academic and disruptive behavior that occurred in a school setting. They included studies that used either a single-subject or group experimental design. Results found that 52.2% of the included studies reported treatment integrity data, which represents an improvement over previously reported levels found in the behavior analytic literature. It is also possible that the increase in reported treatment integrity could be attributed to the journals included in this study rather than an acutal increased in reported treatment integrity. The culmination of these studies together demonstrates that although

there may be an increasing trend in the level of reported treatment integrity, the current levels are still low and at best only 50% of studies are measuring the independent variable. This lack of focus on the independent variables is not only captured in low levels of reported treatment integrity data, but it is also evident in the lack of studies that have manipulated it to understand its effect on treatment outcomes.

Challenges to and Interventions to Improve Treatment Integrity

Not only is it important to report treatment integrity data in research but also for treatment integrity itself to be the focus of research. Research needs to not only address questions of how effective an intervention is when it is implemented but also how effective it is when implemented with low treatment integrity. This type of research is needed to understand how different levels of treatment integrity affect intervention outcome. It is important to evaluate interventions first under ideal circumstances (100% integrity) to establish general utility of the intervention. That is, if an intervention is not effective under the perfect conditions, then it is not likely to be effective under less than perfect conditions. This first step to intervention evaluation is necessary, but is not sufficient. Not understanding how effective specific treatments will be when implemented with less than perfect treatment integrity presents a threat to the practical utilization of the interventions in socially significant contexts.

External validity refers to the extent to which an intervention is generalizable to other settings, populations, and contexts. Interventions that are effective only in analog or research settings have little value to practitioners and parents who will not have the luxury of high control over the setting, interventionist, or context, as is the case with researchers. If treatments are not effective at the levels of integrity that can be sustained in the natural

environment they have little utility outside of the research context and lack external validity, which is a crucial factor in both research and in practice (Peterson et al., 1982; Yeaton & Sechrest, 1981).

Gresham (1989) suggested that factors such as treatment complexity, time required implementing an intervention, materials needed for an intervention, number of treatment agents, perceived or actual effectiveness of an intervention, and motivation of the interventionist may affect treatment integrity. More specifically, the more complex the treatment is the more likely the interventionist is to intentionally or unintentionally fail to implement all parts of an intervention. They may intentionally not implement the intervention as intended because they have too many other demands and need to simplify the intervention to make it manageable. Alternatively, they may fail to implement a treatment with good integrity by forgetting components without being aware that they were not implementing the intervention correctly.

Materials may also play a role in poor treatment integrity in that if the interventionist does not easily have access to needed materials (e.g. reinforcers, self-monitoring sheets), they may supplement with items that are less effective as the originally identified items or simply leave out that component. As the number of treatment agents increases, treatment integrity may decrease due to the intervention requiring coordination between multiple individuals. For example, check-in/check-out (Campbell & Anderson, 2011), a common intervention used to address academic and/or behavior issues in schools often involves both a mentor for the student to check-in/out with every day, the teacher(s) to fill out the daily behavior checklist, and the parent to deliver an external reward at home. This coordination of adults leaves many

opportunities for treatment failures to occur. Also, issues related to the expectation of an intervention and individual differences in motivation to implement the intervention present areas that are also vulnerable to integrity failures.

The factors disused by Gresham (1989) are a rationally derived sample of issues that may contribute to treatment integrity failures and highlight some of the diverse reasons that treatment integrity failures occur. Despite this well articulated list of threats to treatment integrity, relativity little research exists in the literature. These threats to treatment integrity remain theoretical and need to be systematically assessed and manipulated to understand their effects on treatment outcomes (Gresham, Gansle, Noell, Cohen & Rosenblum, 1993). One area that has received some attention has focused on increasing treatment integrity for school-based interventions. Witt et al. (1997) and Noell et al. (1997) both identified performance feedback as an effective means of increasing teacher implementation. The Witt et al. study included four teacher-student dyads. Dyads were included in the study if the nature of the student's problem was performance, not skill, based. During teacher training, teachers learned to implement an intervention designed to target motivation rather than skill acquisition and received all materials needed to conduct the intervention correctly. Following the training phase teachers implemented the intervention independently without feedback. Measurement of treatment integrity occurred via permanent products created at each step of the treatment; these permanent produces served as a means to measure the percentage of steps completed for each teacher. Following the independent implementation phase, a performance feedback phase was conducted. During performance feedback a consultant met with each teacher daily. Daily meetings consisted of graphic display of teachers' current performance, tips on how to

increase implementation, and, for one teacher, individualized training. Following performance feedback, a maintenance phase was conducted which was identical to the independent implementation phase. Treatment integrity was high during training, indicating that the teachers were both informed and skillful about how to correctly implement the intervention. However, when teachers independently implemented the intervention the percentage of steps completed markedly decreased for all for participants. Once performance feedback was initiated the percentage of completed steps increased to 100% and remained high during the maintenance phase for three of the four participants. The results of this study indicated that performance feedback was a powerful and effective procedure for increasing treatment integrity for teachers.

Noell et al. (1997) extended and replicated the work of Witt et al. (1997). They replicated the finding that performance feedback was effective for increasing teacher treatment integrity but extended the work by including measurement of student performance. Their study included 3 regular education elementary teacher-student dyads. The dyads were only included if the referral concern was an academic performance problem. They collected data via permanent products on teacher treatment integrity (percentage of completed steps) and student performance (percentage of correct daily work). The study consisted of three conditions; consultation only (similar to the Witt et al. independent implementation condition), performance feedback, and maintenance. During the consultation only condition, teachers were trained to implement a reinforcement-based intervention and told that someone would collect the data sheets at the end of each day. During the performance feedback condition, consultants met with teachers daily to review a graphic display of current performance with

praise for completed steps and tips for improving treatment integrity. The maintenance condition was the same as the consultation only condition. Results showed that when teachers received consultation only they initially completed high number of steps but treatment integrity quickly fell to low levels. However, once performance feedback was initiated, treatment integrity increased to high levels but only remained high for one teacher during maintenance. One of the two teachers whose performance decreased in maintenance was re-exposed to performance feedback, which led to an immediate increase in performance. These results replicated the findings by Witt et al. indicating that performance feedback is an effective and efficient mechanism for increasing teacher treatment integrity.

In the Noell et al. (1997) study, the results for student performance were idiosyncratic across students, making interpretation difficult. One student's academic performance increased during the consultation only phase and remained high throughout the study. Another student's academic performance increased only after the performance feedback and remained high throughout the remainder of the study and a third student's academic performance did not improve. There are several potential explanations for these idiosyncratic results. For the student who did not improve, it is possible that a more powerful intervention was needed to achieve a positive outcome, therefore it is possible that a different intervention would have produced a more positive outcome. It is also possible that for this participant, exposure to the intervention when implemented with poor integrity influenced the effectiveness of the intervention even once implemented correctly. That is, prior exposure to the treatment with poor integrity may have influenced the overall intervention outcome. For the two participants who did improve academically it is interesting to note that once their academic performance

improved, it remained high despite less than perfect treatment integrity. In addition one student showed improvements during the consultation phase when treatment was poorly implemented meaning that high treatment integrity was not necessary to achieve increases in academic performance for this student. However the other student improved only once integrity was high. These results underscore the need to understand how effective particular interventions are when implemented at varying degrees of integrity. They also highlight the potential for results to be highly idiosyncratic and influenced by individual factors such as history and reinforcement preference.

In a follow up study to the Witt et al. (1997) and Noell et al. (1997) studies, Noell et al. (2005) conducted a study where they both replicated and extended the previous findings on the effectiveness of performance feedback as a tool in increase teacher treatment integrity. Their study included 48 teachers who were predominantly female and taught at six different urban elementary schools. The school populations were almost exclusively African American and had a high level of poverty. Random assignment was used to assign teachers to one of three follow-up conditions. The three conditions included weekly follow-up, which consisted of a brief weekly meeting where teachers were asked about the extent to which they had implemented the intervention. A commitment emphasis condition was similar to the weekly follow-up condition but in addition to a weekly meeting, it included a social influence procedure prior to the intervention implementation. The third condition was performance feedback, which was similar to the performance feedback conditions included in both the Witt et al., and Noell et al. studies. In performance feedback, consultants initially met with teachers daily but that schedule was quickly thinned so that they met with teachers on a weekly basis. During the

meeting consultants reviewed a graphic display of current performance with praise for completed steps and tips for improving treatment integrity

The results from the Noell et al. (2005) study yielded a significant difference between the performance feedback condition and the other two follow-up conditions, indicating that treatment integrity was higher when performance feedback was used compared to the other to follow-up strategies. They also measured student outcomes under each of the follow-up conditions and found that student outcomes were substantially better when performance feedback was implemented as compared to the other follow-up conditions, indicating that there was a link between high treatment integrity and positive student outcomes. The accumulation of these studies provides strong evidence that performance feedback can be an effective method for increasing treatment integrity which is important and valuable information for practitioners. However, because these studies did not evaluate interventions on an individual level, there is still a need to understand more about how effective individual interventions are when they implemented with less than perfect integrity.

Treatment Integrity as a Moderator of Treatment Effect

The need to understand intervention effectiveness under different levels of integrity represents another important aspect of treatment integrity research. By not evaluating interventions under varying levels of treatment integrity, the generalizability of interventions is threatened. Interventions implemented in the natural environment have a greater likelihood of poor treatment implementation due to the complexity of the natural environment and the sustained nature of implementation. Parents and teachers are likely to have multiple completing contingences at any given time. These competing contingences may increase the

likelihood of making an error and, because these errors are likely to occur, it is important to study and understand their effect on specific treatments' effectiveness. To date there are only a handful of studies that have systematically manipulated treatment integrity to understand how effective specific treatments are when components of the intervention are completely omitted or implemented with less than perfect integrity.

Differential reinforcement is a commonly used treatment for both the reduction of problem behaviors and skill acquisition. It has proven to be effective for reducing a variety of maladaptive behaviors (e.g., aggression, self-injury, disruptive behavior) (Deitz & Repp, 1983; Marcus & Vollmer, 1995) and increasing adaptive behaviors (Karsten & Carr, 2009; Roberts, Nelson, & Olson, 1987) in a variety of contexts (e.g., schools, clinics, homes). Differential reinforcement typically involves both a reinforcement component and an extinction component (e.g., no longer providing reinforcement for a previously reinforced behavior or providing escape for problem behavior). Two common types of differential reinforcement include differential reinforcement of other behavior (DRO) and differential reinforcement of alternative behavior (DRA). In a DRO procedure, reinforcement is contingent upon the absence of a target behavior, while in a DRA procedure reinforcement delivery occurs contingent on the occurrence of an alternative behavior. Both differential reinforcement procedures are empirically supported in the behavior analytic literature when implemented with high levels of treatment integrity (Carr & Durand,1985; Vollmer & Iwata, 1992; Vollmer, Iwata, Zarcone, Smith, & Mazalski, 1993;); however less known about the effectiveness of DRO and DRA when they are implemented with less than perfect integrity.

Gresham (1989) highlighted some potential reasons why caregivers and teachers may not implement treatments with high integrity such as the complexity of the natural environment. Factors such as treatment complexity (the more complex, the higher the chance of integrity failure), competing activities (care of other children), history of reinforcement (how long the behavior has been occurring and reinforced), and treatment drift (interventionists slowly changing components of intervention) may also be reasons for decreased treatment integrity. For example, if a teacher implements a DRA which involves both reinforcement and extinction, they may not implement all of the components due to competing activities that are co-occurring in the classroom. This lack of treatment integrity may simply occur to make the classroom demands more manageable for the teacher. The poor treatment integrity may also be due to the student's reinforcement history. Problem behavior that has been reinforced for long periods of time may increase the likelihood of an extinction burst, which may make the extinction component of differential reinforcement difficult or simply impossible to implement with high integrity. An additional implementation challenge is that as time passes the interventionist may change components of the intervention (drift), possibly without being aware that they are no longer following the protocol correctly.

One way these factors have been mitigated against in the past is with thinning procedures. There are several ways in which thinning can occur. Thinning can include systematically removing components of an intervention or thinning of reinforcement schedules. Either way thinning typically occurs over an extended period of time in an effort to maximize treatment effectiveness and minimizing treatment components. Thinning has been demonstrated to be an effective means for simplifying interventions (Hagopian, Fisher, &

Legacy, 1994; Piazza, Moes, & Fisher, 1996; Roane, Fisher, Sgro, Falcomata, & Pabico, 2004) but because it is often a lengthy process it is not always practical in more natural setting. Thinning may be one option for making treatments more user-friendly, but may not be a viable option for practitioners due to the length of the process.

Despite the clear clinical need and numerous calls in the literature to increase focus on treatment integrity, there are still only a handful of studies that have directly investigated this topic. The studies that do exist have found mixed results as to the necessity of high treatment integrity. This inconsistency in the literature may be the result of several factors such as populations, functions of behavior, reinforcement history, and procedural differences in studies.

Extinction is often considered a necessary component in differential reinforcement procedures, however it is also often the most difficult component for interventionist to implement due to extinction bursts and/or the social unacceptability of ignoring problem behavior (e.g, attention extinction). However there is mixed evidence for the necessity of extinction in differential reinforcement procedures. For example, Mazaleski, Iwata, Vollmer, Zarcone & Smith (1993) found extinction necessary to produce positive outcomes. Their investigation included three women diagnosed with intellectual disabilities who engaged in a variety of topographies of self-injurious behavior (SIB). Two of the three participants experienced the DRO procedure both with and without extinction. For these two participants extinction was a necessary component of the DRO procedures even if the reinforcers were highly preferred. The authors of this study concluded that extinction was a “critical factor” in DRO to be an effective intervention.

Although the Mazaleski et al. (1993) study indicated the need for extinction in differential reinforcement procedures in order to be effective, other studies have found that it is not always necessary. For example, Shirley, Iwata, Kahng, Mazaleski and Lerman (1997) investigated the necessity of extinction in DRA procedures using functional communication training (FCT) to treat SIB with three individuals diagnosed with intellectual disabilities. This study included three different conditions: baseline, FCT without extinction, and FCT with extinction. In the baseline condition the functional reinforcer was delivered contingent on problem behavior and no reinforcement was delivered contingent on the FCT response. During the FCT without extinction condition, reinforcement was delivered for both problem behavior and the FCT response. For all three participants, initial exposure to FCT without extinction was not effective in reducing their rate of SIB, but when extinction was included, FCT was effective in reducing rates of SIB. Participants were then exposed to FCT without extinction a second time; during this phase FCT without extinction maintained the same low levels of SIB that were found with extinction for two of the three participants. For the one participant whose SIB increased during the second exposure to FCT without extinction, they were re-exposed to FCT with extinction which successfully reduced SIB. In conclusion, Shirley et al. found that extinction was necessary unless the FCT without extinction followed a phase in which extinction was in place. This study provides some evidence that high treatment integrity may be most important during initial exposure to an intervention.

Athens & Vollmer (2010) also investigated the need for extinction in DRA with six individuals diagnosed with autism and one individual diagnosed with ADHD who all engaged in some form of aggression. Their study manipulated various reinforcement dimensions to

understand the effects on DRA interventions without extinction. They evaluated reinforcement delivery duration, reinforcer quality, and delay to reinforcement in three separate evaluations. For the two participants included in the analysis of duration, they found that when duration of reinforcement for appropriate behavior was either equal to or exceeded the duration of reinforcement for problem behavior, participants engaged in more appropriate behavior. For the two participants that were included in the analysis for reinforcement quality, they found that when reinforcer quality was higher for appropriate behavior then it was for inappropriate behavior the participants engaged in more appropriate behavior. For the two participants that were included in the analysis for reinforcement delay, when reinforcement was delayed 60 s for problem behavior yet delivered immediately for appropriate behavior, the participants engaged in more appropriate behavior. In summary, Athens and Vollmer found that when reinforcement dimensions are manipulated to favor appropriate behavior extinction is not necessary to achieve positive treatment outcomes. The implication of these studies on treatment integrity is that there are many ways in which treatments could be modified so that complete integrity of the extinction component would not be necessary to achieve a therapeutic effect.

Although these studies clarify the role of extinction in differential reinforcement, some questions are unanswered, such as the robustness of interventions when implemented intermittently. There have been several studies that have investigated how intermittent implementation of various components of differential reinforcement affects treatment outcomes. For example, Northup, Fisher, Kahng, Harrell, and Kurtz (1997) examined the effects of intermittent reinforcement and punishment using a DRA treatment. The study included two

children and one adult diagnosed with developmental disabilities. The participants engaged in a variety of maladaptive behaviors including aggression, SIB and pica. The study included a baseline condition in which the functional reinforcer was provided contingent on problem behavior. Following baseline participants were exposed to a 100% treatment condition in which DRA with time-out was implemented with 100% integrity. Following this phase, participants were exposed to a variety of conditions in which either the reinforcement or punishment or a combination of reinforcement and punishment were delivered with 50% or 25% integrity. They found that the interventions remained effective even when implemented with 50% integrity and that only slight increases in problem behavior emerged when implemented at 25% integrity. These results provided preliminary evidence that differential reinforcement may be a robust treatment which is effective even when treatment integrity is relativity low. However, because the study included punishment as a treatment component, it is unknown if DRA without punishment would be as robust.

Vollmer, Roane, Ringdahl, and Marcus (1999) extended the results found by Northup et al. (1997) by evaluating two types of treatment integrity errors in DRA. They evaluated both omission errors and commission errors. Omission errors are leaving a component out whereas commission errors are adding a component to the intervention. The commission errors consisted of delivering reinforcement for problem behavior (i.e., not implementing the extinction component with 100% integrity). Omission errors consisted of not delivering reinforcement for appropriate behavior. They found that DRA was effective even when 50% of the intervals involved either a commission or omission error. These results support the findings of Northup et al. that suggest that DRA is a robust treatment even when implemented with less

than perfect integrity. However, since all the participants experienced the treatment with 100% integrity prior to treatment integrity failures, it is possible that exposure to the intervention at 100% integrity influenced the robustness of the intervention. It is possible that had the treatment been implemented first with low levels of integrity that the treatment would have produced outcomes that were less favorable. As the data are, Vollmer et al. extended the work of Northup el al. by evaluating DRA without the use of extinction and evaluating the effects of two types of errors.

St. Peter-Pipkin, Vollmer, & Sloman (2010) sought to replicate and extend the literature on the extent to which commission errors and omission errors affect treatment outcomes of DRA treatments using human operant procedures with one applied replication. They evaluated commission and omission errors both in isolation and in combination with each other. In addition, they sought to understand if exposure to DRA without errors prior to exposure to DRA with errors influenced the DRA outcomes. The results of the human operant experiment verified previous research. When omission errors occurred in isolation, DRA remained effective at 100%, 80%, and 60% treatment integrity. However, when treatment integrity fell to 40% and 20%, participants engaged in slightly lower levels of responding. When commission errors occurred in isolation and/or in combination during 40% or 20% of intervals, the DRA procedure was markedly less effective. This was especially true if the low integrity DRA phase followed a baseline phase. These results replicated previous findings that DRA is a robust intervention that is effective even when implemented with less than perfect integrity. However if treatment integrity falls below 50%, treatment outcomes are compromised. In addition, this study found that exposure to DRA at full integrity prior to treatment integrity failures can insulate and

maintain the treatment effects. Although these findings need replication and extension, they highlight the potential need for 100% treatment integrity during initial exposure to DRA interventions.

Another variable that may potentially affect treatment integrity failures is the function of the problem behavior. Function refers to the environmental variables that evoke and maintaining a behavior. The current gold standard for identifying function is the functional analysis technology developed by Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994). Functional analysis research has identified three general reinforcement contingencies for problem behavior: (a) social-positive reinforcement (generally in the form of access to attention and/or tangibles) (b) social-negative reinforcement (generally in the form of escape from instructions or demands), and (c) automatic reinforcement (generally in the form of self-stimulatory behavior) (Carr & Durand, 1985; Derby et al., 2000; Hanley, Iwata, & McCord, 2003).

The most common function of problem behavior identified for the participants in Iwata, Pace, Dorsey et al. (1994) was negative reinforcement (e.g., problem behavior to escape from academic demands). Treatments designed to target problem behavior maintained by negative reinforcement often include prompting as part of the treatment package. Prompting is a mechanism used to prevent escape from the required task or demand. Escape extinction is the continuation of an instructional sequence despite the occurrence of problem behavior and is often a necessary component to effectively treat problem behavior maintained by negative reinforcement (e.g., Iwata, Pace, Dorsey et al.,1994; Kuhn, DeLeon, Fisher, & Wilke, 1999; Lalli, Casey, Goh, & Merlino, 1994).

Although interventions that include escape extinction to treat problem behavior maintained by escape have proven to be effective, escape extinction is often associated with unpleasant side effects (e.g., extinction burst, spontaneous recovery, aggression; Lerman & Iwata, 1995; 1996) and may be difficult or impossible to implement under certain circumstances. Extinction bursts are described as a transitive increase in duration, frequency and/or intensity of problem behavior (Cooper, Heron, & Heward, 2007). Due to the difficulties in correctly implementing escape extinction, treatments that include extinction as a component may be more susceptible to treatment integrity failures. For example, a child may engage in problem behavior at school to escape from doing their work and the prescribed treatment could be DRA for compliance with extinction. The DRA component would consist of a small edible item and the extinction component would consist of a prompting procedure that does not allow for escape. The teacher may implement the treatment at first with 100% integrity, but if the extinction component leads to a temporary increase in the problem behavior, this increase in problem behavior may lead the teacher to not implement that component of the treatment at all or only intermittently.

Prompting is a commonly used tool to teach new skills as well as to ensure correct implementation of escape extinction. The literature to date has investigated the use of a variety of different prompting procedures that have proven to be effective for both acquisitions of new skills as well as ensuring implementation of escape extinction. These prompting procedures can generally be categorized as either least-to-most or most-to-least prompting procedures. Most-to-least prompting procedures implement the “most” intrusive prompt first (e.g., physical guidance) and then systematically move to lesser intrusive prompts (e.g., vocal prompt) based

on a lack of or incorrect response. On the other hand, least-to-most prompting procedures implement the least intrusive prompt first (e.g., vocal prompt) and move to more and more intrusive prompts (e.g., physical guidance) based on a lack of or incorrect response. Prompting procedures have generally been found to be effective (Carbone, Sweeney-Kerwin, Attanasio,& Kasper, 2010; Falcomata, Ringdahl, Christensen, & Boelter, 2010; Yilmaz, Konukman, Birkan, & Yanardag, 2010; Walker, 2008; Bryan, & Gast, 2000) and are commonly included in behavioral interventions that target academics and/or problem behavior. Because prompting is often used to implement the extinction component, it may be a component that is likely to be implemented with less than perfect treatment integrity; therefore it is important to understand how failures to implement prompting contribute to treatment outcomes.

To date there have only been a few studies that have evaluated treatment integrity failures in prompting. For example, Holcombe, Wolery, and Synder (1994) investigated the effects of implementing a constant time delay prompting procedure at two different levels of integrity. Constant time delay is a prompting procedure that is similar to progressive time delay (Touchette, 1971) and typically involves two types of trials. In one trial, the stimulus is presented immediately following the prompt to do work, and in the other (the delay trial), the stimulus is presented after a short fixed-duration delay. During the delay trials, if the child does not respond, prompting occurs to ensure the child completes the task (i.e., does not escape).

The Holcombe, et al. (1994) study included six preschool-aged participants diagnosed with Down Syndrome, Chromosomal Abnormality, or Developmental Delays who were taught to identify pictures. The study included two conditions: a high-fidelity condition and a low-fidelity condition. In the high-fidelity condition, the constant-time delay procedure was implemented

with 100% integrity; in the low-fidelity condition, the controlling prompt was implemented during approximately 50% of trails. The high-fidelity condition was effective for teaching picture identification to five of the six participants and the low-fidelity condition was effective for teaching picture identification to four of the six participants. Although the low-fidelity condition was effective for four of the participants it was less effective (i.e., took longer for them to reach mastery criteria) for three of the participants and was equally effective for only one of the participants. For the one participant for whom only the high-fidelity condition was effective, the authors attempted to teach the un-mastered task using the high-fidelity procedure subsequent to the low integrity procedure; however, the intervention was only effective after they provided reinforcement for each trial. The authors speculated that exposure to the low-fidelity condition may have interfered with learning (i.e., a history of unsuccessful learning). One participant did not master the task when both high and low-fidelity conditions were altered even when two separate tasks were used. For this participant, she only mastered the task when only the high-fidelity condition was used. It is possible for this participant, that exposure to the low-fidelity condition, even when a different task is targeted, may have resulted in her becoming less responsive to the high-fidelity condition.

This study has several important implications. One implication is that prompting failures may result in slower acquisition of learning which could be particularly relevant for fragile populations (e.g., children with developmental disabilities, behavior problems, or slow learners). For these populations, any amount of lost time can be crucial to long-term outcomes. A second implication is that the mere exposure to prompting failures may either prevent or delay learning even after failures are no longer occurring. For example, if a teacher implements

an intervention with poor integrity at the beginning, but later implements the intervention correctly, the exposure to the poorly implemented intervention may affect long-term treatment outcomes despite later corrections. These potential implications may increase the importance of perfect intervention implementation at the beginning of an intervention; however, these findings were only true of one of the six participants included in the study. Therefore, future research would need to replicate these findings.

A second study that evaluated treatment integrity failures in prompting is a study conducted by Noell, Gresham & Gansle (2002). They investigated the effects of implementing instructional prompts at three different levels of integrity. Instructional prompts are vocal prompts that include strategies to complete the presented task. Their study included 6 second-grade children (5 boys, 1 girl) identified as struggling in mathematics by their classroom teachers. The intervention included was a computerized mathematics program that consisted of instructional prompts, corrective feedback, and rewards for correct responses (graphic animations). They included three conditions, a high, medium, and low-fidelity condition. In the high-fidelity condition, no errors in prompting occurred, in the medium-fidelity condition, prompting occurred in two-thirds of trails, and in the low-fidelity condition, prompting only occurred in one-third of trails.

In general, the high-fidelity condition was the most effective, followed by the medium-fidelity condition. The low-fidelity condition in which the prompts were only presented one-third of the time was the least effective, but the differences between the medium and low-fidelity conditions was modest at best and idiosyncratic both across participants and conditions. These results replicated the work done by Holcombe, et al. (1994) in that decreases in

prompting fidelity lead to decreases in treatment effectiveness; however, these studies leave many questions still unanswered such as specific levels of treatment integrity needed for treatment to remain effective, which treatments are more robust, and the effects of learning histories.

An additional question in need of investigation is the effect of prompting failures on treatments that rely on prompting to implement escape extinction. Most of the research to date has investigated prompting failure effects on skill acquisition and has not focused on these failures in relation to extinction. It is unknown if prompting failures would lead to a decrease in intervention effectiveness or if interventions would remain effective despite lower levels of treatment integrity for prompting.

Another aspect of treatment integrity was investigated by Gansle and McMahon (1997). They sought to understand the extent to which failures to implement components of a self-monitoring program in a classroom affected the outcome of the intervention. They found that self-monitoring was a robust treatment that remained effective even when graphing and reward components were not implemented with perfect integrity. The study included 21 3rd through 6th grade public school teachers and 49 students, each assigned to one of three treatment conditions. In the 100% integrity condition all components of the self-monitoring intervention were implemented which included self-monitoring with feedback, reward, and graphing; in the 83.3% integrity condition self-monitoring with feedback and reward were implemented but the graphing component was not included; in the 66.7% integrity condition only self-monitoring was implemented. The results indicated that higher levels of treatment integrity were not predictive of student outcome, meaning that the reward and graphing

components may not be essential components of an effective self-monitoring intervention. It is possible that these results are limited to this population; therefore, these results need replication. However, it is important to note that this study highlights the need for research to be conducted on treatment integrity, because there may be a variety of interventions for which only certain components are essential. In other words, the complexity associated with many common interventions may not be necessary to achieve the desired outcomes and simplification of the intervention may lead to better treatment integrity in the natural environment.

Although the number of studies focused on treatment integrity has increased there are still many questions left unanswered. One such question is the role function may play in treatment integrity. It is clear from the literature that function of problem behavior has a key role in treatment identification (Iwata, Pace, Cowdery et al., 1994) and this role of function has played a pivotal role in both research and practice. For example, the Individuals with Disabilities Act (IDEA), (1997), requires the inclusion of an FBA if an individual with an Individualized Education Plan (IEP) is subject to disciplinary action. This increase in the use of function-based interventions in schools and practice means that we need to understand more about how effective these interventions are when conducted in these environments. It is possible that treatment integrity failures will differentially affect treatment outcomes depending upon the function of the problem behavior, much like treatments can be differentially effective depending upon the problem behavior. For example, time-out, a commonly used intervention both in the home and school setting, is not likely to be effective for an individual whose problem behavior is maintained by escape. This is because time-out is a form of escape

provided for problem behavior. In this situation time-out may actually increase problem behavior since the functional reinforcer is contingent on problem behavior. On the other hand, time-out is likely to be highly effective for an individual whose problem behavior is maintained by attention. For this individual the functional reinforcer is removed contingent upon the problem behavior which will likely lead to decreases in problem behavior. In this example the same treatment has different outcomes depending upon the function of the problem behavior.

Much like treatments having differential outcomes depending upon function it is also possible that treatment integrity failures would differentially affect treatment outcomes depending upon the function. For example, a token economy, which might consist of delivering a penny with praise contingent on each task completed and a prompting component to prevent escape may be used in a classroom to increase on-task behavior. If the problem behavior was determined to be maintained by escape, it is possible that failures in the prompting component could be more detrimental than failures in reinforcer delivery for this individual. On the other hand, if the individual's problem behavior is maintained by attention, failures to deliver the penny with praise may be more detrimental to treatment outcomes than failures in prompting. For the individual whose problem behavior is maintained by escape, the prompting component is preventing the functional reinforcer from being delivered contingent on problem behavior. On the other hand, for the individual whose problem behavior is maintained by attention, the functional reinforcer of praise and a penny are delivered contingent on appropriate behavior. However, at this point, the effects of function on treatment integrity remain largely unknown and, due to increasing importance of function-based treatments, these relationships need investigation.

Although the research on treatment integrity is growing, it appears to be growing at a slow and uneven rate. For example, there has been a large and extensive focus on the lack of reported treatment integrity in the behavioral literature, yet despite this clear need and strong evidence for increasing reporting of treatment integrity, there has been only modest improvement since the publication of the Perterson et al. (1982) study. These repeated calls to increase reporting of treatment integrity should have lead to an increase of reported treatment integrity and an increase in the study of treatment integrity, but it has not. One area that is clearly very important to the concept of treatment integrity is how to get adults to simply implement interventions. This area has also received a significant amount of focus and led to the development of performance feedback which has proven to be an effective strategy for increasing levels of implementation (Noell et al., 2005). An area that remains in critical need of examination is the focus on understanding the role of treatment integrity in commonly used interventions. When you consider the multitude of available interventions, only a handful of these interventions have been studied in regards to treatment integrity. This apparent lack of literature in this area leaves many questions still unanswered.

This leads to the purpose of the current investigation which is two-fold. First, to understand the effect of prompting failures and reinforcement failures for differential reinforcement procedures used to increase academic compliance. This question was answered by exposing participants to a full-treatment condition, in which both prompting and reinforcement are implemented, and two half-treatment conditions, in which participants were exposed to one condition where only reinforcement was available and one in which only prompting was available. The hypothesis is that treatment integrity failures of reinforcement

will be more detrimental to treatment outcomes as opposed to treatment failures of prompting. An additional purpose is to understand the role that preferences for either escape or attention play in differentially affecting treatment integrity failures in prompting and reinforcement delivery. The hypothesis is that individuals who prefer attention as a reinforcer will be more sensitive to failures in reinforcement than failures in prompting and individuals who prefer escape as a reinforcer will be more sensitive to failures in prompting than failures in reinforcement.

METHOD

Participants and Setting

Participants included 36 children who were enrolled in either kindergarten or first grade ranging in age from 5 to 7 years old. The ethnic make-up of the participants was approximately 53% African American, 28% Caucasian, 11% mixed-ethnic descent and 8% Asian. Forty-four percent were male and 56% female. Students were recruited from either a small private school or from one of two public urban elementary schools in southern Louisiana. A parental consent form was sent home to all kindergarten and first grade classrooms in each of the participating schools. Students who returned consent forms and gave assent were included in the study. Prior to conducting the study sample size was determined using the computer program G-Power 3.1, which indicated that 15 participants were needed to achieve adequate power. The power level was set at 0.95 with a modest effect size of 0.4. All sessions were conducted either in a small room made available by the school or in a corner of the school library.

Data collection

Data were collected on the following measures: independent correct responses, independent incorrect responses, prompted correct responses and prompted incorrect responses. An independent correct response was defined as initiating or completing one item on the worksheet correctly within 5 s of start of the session or within 5 s of completing a previous item. An independent incorrect response was defined as initiating or completing one item on the worksheet incorrectly within 5 s of the start of the session or within 5 s of completing a previous item. A prompted correct response was defined as completing or

initiating one item correctly after a prompt was provided. A prompted incorrect response was defined as completing or initiating one item incorrectly after a prompt was provided.

Treatment Integrity: All sessions were conducted by the first author who had extensive experience conducting behavioral interventions similar to the intervention included in this study. Treatment integrity data were collected on correct delivery of positive reinforcement, incorrect delivery of positive reinforcement, correct prompting and incorrect prompting. Correct delivery of positive reinforcement was defined as delivering the reinforcer within 2 s of a correct response. Incorrect delivery of positive reinforcement was defined as the omission of reinforcer delivery within 2 s of a correct response. Correct prompt delivery was defined as presenting the prompt after 5 s of no work. Incorrect prompt delivery was defined as omission of a prompt after 5 s of no work. To calculate treatment integrity, the number of correct responses for both prompt delivery and reinforcer delivery was divided by the number of correct responses plus the number of incorrect responses and multiplied by 100%. Treatment integrity data were collected on 100% of sessions. Mean treatment integrity was 99.33% (range, 50%-100%).

Interobserver Agreement: HP minicomputers were used to collect data using real time data collection software. To assess interobserver agreement, a second observer collected data on 25% of sessions on the outcome measures. Data collectors were graduate students at Louisiana State University and were trained prior to taking data for the project. Each condition was 5 min in duration and for each participant all sessions were conducted on the same day. The primary author was present in the room at all times.

Observers' records were divided into 10-s intervals for the purpose of calculating observer-agreement coefficients. Each record was compared on an interval-by-interval basis using a proportional agreement method. That is, each interval scored with the same frequency of target response was scored as 1, each interval not in exact agreement was provided a proportional agreement score by dividing the smaller measure by the larger measure. The score for each interval was then summed then divided by the total number of intervals, and converted into a percentage agreement. The mean IOA for independent correct responses, independent incorrect responses, prompted correct responses and prompted incorrect responses were 84.60%(range 19.35% to 100%), 96.06% (range, 77.42% to 100%), 98.81% (range,86.56% to 100%) and 97.23 %(range, 22.58% to100%) respectively. The mean IOA for correct delivery of positive reinforcement, incorrect delivery of positive reinforcement, correct prompting and incorrect prompting were 93.09% (range, 55.11% to 100%), 99.69% (range, 93.55% to100%), 98.73% (range, 82.26% to 100%) and 100%, respectively.

Procedures

Reinforcer Assessment. Prior to the experiment, participants' preferences for either attention or escape were assessed directly using procedures similar to St. Peter Pipken, et al. (2010). At the onset of the assessment, participants were exposed to a forced-exposure trial. During the forced-exposure trials two math worksheets were placed in front of them that were identical except one worksheet had the word "Break" written on top and one worksheet had the word "Talk" written on top. They were then given the following verbal instructions:

> You have two worksheets in front of you that are exactly the same expect one has the word break written on top and one has the word talk written on top. If you choose the break worksheet then for every problem you complete I will let you have a short break.

If you choose the talk worksheet, for every problem you complete I will tell you what a great job you are doing.

They were then prompted to touch the "break" worksheet. If they did not touch the worksheet within 5 s, the experimenter physically guided them to touch the worksheet and prompted them to complete one item every 5 s by saying "Do your work." During the forced-exposure trials prompting continued every 5 s until they completed one item. Once they completed the item, both worksheets were removed for 10 s. The participants were then prompted to touch the "talk" worksheet. If they did not touch the worksheet within 5 s, the experimenter physically guided them to touch the worksheet. If they did not complete an item within 5 s the experimenter physically guided them to complete an item. Once the item was complete the experimenter provided them with brief eye contact and descriptive praise (e.g., "you did a great job on that problem," "Wow, you sure are smart," "I love the way you are doing your work").

Following the forced-exposure trials they were given the following verbal instructions: You have two worksheets in front of you that are exactly the same except one has the word break written on top and one has the word talk written on top. If you choose the break worksheet then for every problem you complete I will let you have a short break. If you pick to the talk worksheet, for every problem you complete I will tell you what a good job you are doing. You can switch worksheets whenever you like or you can choose to do nothing.

Following the instructions the same two worksheets used in the forced-exposure trials were placed in front of them. Contingent on choosing (either verbally or physically by touching the worksheet) the worksheet that had the word "Break" written on it, they were prompted to

complete an item by saying “Do your work” every 5 s until they completed one item on the worksheet. Once they completed one item, both worksheets were removed for 10 s. After 10 s both worksheets were placed in front of the participant. Contingent on choosing the “Talk” worksheet, the graduate student did not make eye contact or talk to participants until they completed one item. Once they completed one item they were provided with eye contact and brief verbal praise, such as: “you did a great job on that problem,” “Wow, you sure are smart,” “I love the way you are doing your work.” These procedures were repeated for 2 min. The purpose of this assessment was to determine participants’ preference for either attention and/or escape as a reinforcer for this academic task.

Experimental Analysis. Directly following the reinforcer assessment participants were exposed to four different conditions: *Baseline, Full Treatment, Reinforcer Only, and Prompting Only*. Each condition was 5 min in duration. The order of conditions was counterbalanced across participants and was determined by creating a spreadsheet with all possible sequences. Participants were randomly assignment to one of the 24 sequences such that for the first 24 participants no sequences were repeated. Once all sequences had been assigned the remaining 12 participants were randomly assigned to one of the 24 sequences such that no sequences were repeated was more than twice. Participants in kindergarten were given math worksheets that included counting and matching shapes. Participants in 1st grade were given math worksheets that included single digit addition. The same set of math worksheets were used across all conditions. In addition to the worksheets participants also had access to an alternative distracter item, a portable DVD player. When the children entered the room they were presented with an array of children’s movies and told to choose one movie to watch.

After the participant choose a movie they were shown the reinforcers that could be earned as part of the token economy. The reinforcers consisted of a variety of edible items, small toys and stickers. Items were divided into groups and labeled to indicate how many pennies each item cost. Each participant was given one penny and told "I gave you one penny, which of these items could you get with one penny?" If the child responded correctly they were told "good job that is right." If they did not respond correctly they were told "that is not right, these items over here are the items you can choose from." The participants were then given five pennies and told "Now I gave you five pennies, which of items could you get with five pennies?" If the child responded correctly they were told "good job, that is right." If they did not respond correctly were told "that is not right, these items over here are the items you can choose from." This sequence was repeated using a different number of pennies until the participant responded correctly across two consecutive trials.

At the onset of the *Baseline* condition, each participant was told,

> Here I have some worksheets for you to complete, you can complete them if you want and here I have the movie you chose and you can watch the movie if you want. You can either do the worksheets or you can watch the movie but you cannot do them both at the same time. Also, you can switch between them as often as you like.

There were no programmed consequences for completing items on the worksheet. At the onset of the *Full Treatment* condition participants were told,

> Here I have some worksheets for you to complete, you can complete them if you want and here I have the movie you chose and you can watch the movie if you want. You can either do the worksheets or you can watch the movie but you cannot do them both at the same time, but you can switch between them as often as you like. However, for every item you complete on the worksheet you will get a penny. At the end of 5 minutes, you can exchange your pennies for one of the prizes I showed you earlier. Remember, the more pennies you earn the bigger your prize will be.

The researcher sat next to the participant and prompted compliance every 5 s using a verbal prompt “do your work”, as long as they were not working. Contingent on completing each item they were given one penny and brief praise such as “good job working.” At the end of the session they were allowed to exchange all their pennies for a prize. At the onset of the *Reinforcement Only* condition participants were told,

> Here I have some worksheets for you to complete, you can complete them if you want and here I have the movie you chose and you can watch the movie if you want. You can either do the worksheets or you can watch the movie but you cannot do them both at the same time, but you can switch between them as often as you like. However, for every item you complete on the worksheet you will get a penny. At the end of 5 minutes, you can exchange your pennies for one of the prizes I showed you earlier. Remember, the more pennies you earn the bigger your prize will be.

The condition was identical to the full treatment condition with the following exception; no prompting was given during the session. At the onset of the *Prompting Only* condition the participants were told,

> Here I have some worksheets for you to complete, you can complete them if you want and here I have the movie you chose and you can watch the movie if you want. You can either do the worksheets or you can watch the movie but you cannot do them both at the same time but you can switch between them as often as you like.

These sessions were identical to the full treatment condition with the following exception; the participant did not earn pennies for completing items on the worksheet.

RESULTS

To analyze the primary outcome measure of change within individuals as a result of each intervention (Baseline, Full Treatment, Prompting Only , Reinforcement Only) a repeated measures analysis of variance (ANOVA) was conducted using total correct math problems completed as the dependent variable. This analysis consisted of four levels, one for each treatment condition. To analyze the degree to which preferences for either escape or attention predicted treatment condition effectiveness a repeated measures ANOVA was conducted using total correct math problem completion as the dependent variable and group membership, based on pre-test scores added, as a second factor to the analysis. Follow-up t-tests were conducted to evaluate carryover effects.

Testing Statistical Assumptions

For each t-test statistical analysis, homogeneity-of-variance was tested. If groups were homogenous, equal variance was assumed; however, if the equality-of-variance assumption was violated, results were reported using the equal variance not assumed procedure. Furthermore, for ANOVAs, the assumption of sphericity was tested. If sphericity was violated, the Greenhouse-Geisser correction was used to analyze results.

Treatment Effectiveness

One purpose of this study was to determine the effect of each treatment condition on the impact of the number of correct math problems completed. It was hypothesized that the full treatment condition would yield the greatest increase in correct math completion, followed by reinforcement only, prompting only, and baseline, respectively. Data were collected on correct and incorrect responding throughout each condition. Condition order was randomly

assigned to each participant. To understand the impact of each treatment condition on correct completion of math problems a one-way repeated measures ANOVA was conducted with treatment condition as the repeated factor and the dependent variable being correct completion of math problems. The means and standard deviations for correct completion of math problems for each of the treatment conditions are presented in Table 1. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(5)=12.58, p<.05$); therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε=.78). The results of the repeated measures ANOVA indicated a statistically significant difference between treatments, F (2.35, 33.65) =11.80, p <.01, multivariate η^2=.25. Eta squared was used to calculate the effect size. An η^2=.25 is classified as a large effect by Cohen, (1988). Post-hoc comparisons were conducted using Tukey HSD test, which revealed that the mean score for the baseline condition (M=19.83, SD= 15.31) was significantly different from the full treatment condition (M=31.06, SD= 15.15), reinforcement only condition (M=33.47, SD= 17.84) and prompting only condition (M=31.78, SD= 12.63). No significant differences were reveled between any other conditions

Table 1

Means and Standard Deviations for Correct Completion of Math Problems for Overall Treatment Effectiveness

Treatment Condition	N	*M*	*SD*
Baseline	36	20.40[1]	15.77
Full Treatment	36	33.71[1]	15.23
Reinforcement Only	36	33.20[1]	18.03
Prompting Only	36	31.40[1]	12.61

Note. 1 = significant difference between baseline and full treatment, reinforcement only and prompting only condition $p < .05$ level.

Reinforcer Assessment

A second purpose of the study was to understand the role that preferences for either escape or attention had in predicting which treatment condition would be most effective. It was hypothesized that children who preferred escape would complete more correct problems in the prompting only condition versus the reinforcement only condition, whereas children who preferred attention would complete more correct problems in the reinforcement only condition versus the prompting only condition. To evaluate the treatment effects by group, a repeated measures ANOVA was conducted by group. Three groups of 12 were created based on the percentage of problems completed correctly on the "talk" worksheet compared to the "break" worksheet. Groups were created to evaluate if individuals who responded similarly on the reinforcer assessment would differentially respond to each of the treatment conditions. This allowed for individuals whose responses were similar to be grouped together so that their responses as a group could be evaluated in each condition. Three distinct groups allowed for the creation of a group who preferred escape, a group who preferred attention and a group whose preference was less clear. Groups were created by taking the total number of correct problems completed on the "talk" worksheet and dividing it by the total number of correct problems completed on both worksheets and multiplying it by 100. Participants who completed 50% or less correct math problems on the "talk" worksheet were assigned to the break group, participants who completed between 51% and 89% of correct problems on the "talk" worksheet were assigned to the ambiguous group and participants who completed more than 89% of correct problems on the "talk" worksheet were assigned to the attention group. The means and standard deviations of correct completion of math problems for each of the

treatment conditions are presented in Table 2. These data were analyzed by conducting a repeated measures ANOVA for each group. Post-hoc test were conducted to further understand the relationship between each treatment condition and the group.

Table 2

Means and Standard Deviations for Correct Completion of Math Problem by Group and Condition

Group	Treatment Condition	N	*M*	*SD*
Break Group	Baseline	12	11.67[1]	14.98
Break Group	Full Treatment	12	32.08[1,2]	14.63
Break Group	Prompting Only	12	24.75[1,]	11.89
Break Group	Reinforcement Only	12	25.17[1,2]	11.89
Ambiguous Group	Baseline	12	18.92[3]	12.72
Ambiguous Group	Full Treatment	12	35.17[3]	19.49
Ambiguous Group	Prompting Only	12	34.75[3]	13.14
Ambiguous Group	Reinforcement Only	12	37.83[3]	19.97
Attention Group	Baseline	12	28.92	15.98
Attention Group	Full Treatment	12	34.92	15.15
Attention Group	Prompting Only	12	35.83	10.59
Attention Group	Reinforcement Only	12	33.47	19.34

Note. 1 = within the break group a significant difference between baseline and full treatment, reinforcement only and prompting only condition $p < .05$ level. 2= within the break group a significant difference between full treatment and reinforcement only condition $p < .05$ level. 3= within the ambiguous group a significant difference between baseline and full treatment, reinforcement only and prompting only condition $p < .05$ level.

Break Group. For the break group (completed ≤50% of problems on the "talk" worksheet") Mauchly's test indicated that the assumption of sphericity had not been violated ($\chi^2(5)=7.00, p>.05$). Means and standard deviations for the break group are found in Table 2. The results of the one-way repeated measures ANOVA indicated a statistically significant difference between conditions for the break group, F (3, 9) =5.84, p =.003, multivariate η^2=.35. The obtained η^2 is classified as a large effect by Cohen (1988). Post-hoc comparisons were conducted using Tukey HSD test which revealed that the mean score for the baseline condition

(*M*=11.37, *SD*= 14.98) was significantly different from the full treatment condition (*M*=32.08, *SD*= 14.63), reinforcement only condition (*M*=25.17, *SD*= 11.26) and prompting only condition (*M*=24.75, *SD*= 11.89); indicating that in all three treatment conditions participants in the break group completed significantly more problems than were completed in the baseline condition. A significant difference was also found between the full treatment condition (*M*=32.08, *SD*= 14.63) and reinforcement only (*M*=25.17, *SD*= 11.26) condition; indicating that the break group participants completed significantly more problems in the full treatment condition when compared to the reinforcement only condition. No other significant differences were reveled between any other conditions.

Ambiguous Group. For the ambiguous group (completed 50% to 89% of problems on the "talk" worksheet) Mauchly's test indicated that the assumption of sphericity had not been violated ($\chi^2(5)=9.43, p>.05$). Means and standard deviations for the ambiguous group are found in Table 2. The results of the one-way repeated measures ANOVA indicated a statistically significant difference between conditions for the break group, F (3, 9) =6.38, p =.002, multivariate η^2=.37. Eta squared indicated a large effect. Post-hoc comparisons were conducted using Tukey HSD test which revealed that the mean score for the baseline condition (*M*=18.92, *SD*= 12.72) was significantly different from the full treatment condition (*M*=35.17, *SD*= 19.49), reinforcement only condition (*M*=37.83, *SD*= 19.98) and prompting only condition (*M*=34.75, *SD*= 13.14); indicating that in all three treatment conditions participants in the ambiguous group completed significantly more problems than were completed than in the baseline condition. No other significant differences were found between any other conditions.

Attention Group. For the attention group (≥90% of problems on the "talk" worksheet) Mauchly's test indicated that the assumption of sphericity had not been violated ($\chi^2(5)=6.94, p>.05$). Means and standard deviations for the attention group are found in Table 2. The results of the one-way repeated measures ANOVA indicated no statistically significant difference between the conditions, F (3, 9) =1.39, p >.05; therefore no post-hoc analyses were conducted for the attention group.

Condition Carryover Effects.

In an effort to understand the role of condition order on participant responding, follow-up t-tests were conducted to analyze these effects. T-tests were conducted to further understand the influence of exposure to reinforcement on subsequent conditions. Independent t-tests evaluated (a) whether exposure to a reinforcement condition (full treatment or reinforcement only) prior to baseline significantly influenced the number of problems completed in baseline, (b) if the first exposure to a reinforcement condition influenced the number of problems completed in the second exposure to a reinforcement condition. In addition, two independent t-tests were conducted to further understand the influence of exposure to prompting on subsequent conditions. Independent t-tests evaluated (a) if exposure to a prompting condition (full Treatment or prompting only) prior to baseline significantly influenced the number of problems completed in baseline, (b) if the first exposure to a prompting condition influenced the number of problems completed in the second exposure to a prompting condition.

To evaluate the role of exposure to reinforcement on baseline a paired-sample t-test was conducted comparing the mean of problem completion when baseline was and was not

preceded by a reinforcement condition (full treatment or reinforcement only). On average, participants who experienced the baseline condition prior to a condition in which reinforcement was available (*M*=8.3, *SE*=4.32) exhibited lower baseline scores (*M*=24.27, *SE*=2.87, *t*(34)=-2.99, *p*<.05,*r*=.46)than those who experienced a reinforcement condition prior to baseline. Means and standard deviations are found in Table 3. These results indicate that exposure to reinforcement influenced responding in the subsequent baseline condition.

To evaluate the role of prior exposure to reinforcement on subsequent exposure to a reinforcement condition a paired sample independent t-test was conducted comparing the means of the first reinforcement condition to the means of the second reinforcement condition. On average, participants' second exposure to reinforcement yielded significantly higher performance (*M*=35.97, *SE*=3.04) than their first exposure to reinforcement (*M*=31.83, *SE*=2.39, *t* (35) = -2.01, *p*<.05,*r*=.73). Means and standard deviations are found in Table 4. These results indicate that exposure to reinforcement at one time influenced responding when in the participant was re-exposed to reinforcement.

Table 3

Means and Standard Error for Correct Completion of Math Problems in Baseline with and without Prior Exposure to Reinforcement and Prompting

Treatment Condition	N	*M*	*SE*
Baseline (No Reinforcement Prior)	10	8.30*	4.32
Baseline (Reinforcement Prior)	26	24.27*	2.87
Baseline (No Prompt Prior)	9	9.22*	4.71
Baseline (Prompt Prior)	27	23.37*	2.90

Note. * = significant difference between at the $p < .05$ level.

To evaluate the role of exposure to prompting on baseline an independent t-test was conducted comparing the means of problem completion when baseline was and was not

preceded by a prompting condition (full treatment or prompting only). On average, participants who experienced the baseline condition prior to a condition in which prompting was available (*M*=9.22, *SE*=4.71) had baseline scores that were significantly lower (*M*=23.37, *SE*=2.90, *t*(34)= -2.47,*p*>.05,*r*=.39) than those who experienced a prompting condition prior to baseline. Means and standard deviations are found in Table 3. These results indicate that exposure to prompting did influence responding in subsequent baseline conditions.

To evaluate the role of prior exposure to prompting on subsequent exposure to a prompting condition, a paired sample t-test was conducted comparing the means of the first prompting condition to the means of the second prompting condition. On average, participants' second exposure to prompting was not significantly higher (*M*=33.17, *SE*=2.11) than their first exposure (*M*=32.67, *SE*=2.53, *t*(35)=2.13,*p*>.05,*r*=.50). Means and standard deviations are found in Table 4. These results indicate that exposure to prompting at one time did not influence responding when the participant was re-exposed to prompting.

Table 4

Means and Standard Error for Correct Completion of Math Problems in First and Second Reinforcement and Prompting Conditions

Treatment Condition	N	*M*	*SE*
First Reinforcement Condition	36	31.83*	2.39
Second Reinforcement Condition	36	35.97*	3.04
First Prompt Condition	36	32.67	2.53
Second Prompt Condition	36	33.17	2.11

Note. * = significant difference between at the $p < .05$ level.

To further evaluate carryover effects a paired-sample t-test was conducted comparing the means of the first condition to the means of the last condition. On average, participants correct responses in the first condition (*M*=24.62, *SE*=2.80) were statistically lower than their

responses in their last condition (*M*=35.06, *SE*=2.70, *t*(35)=-3.16,*p*>.05,*r*=.47). Means and standard deviations are found in Table 5. These results provide additional evidence that order of condition may have influenced the results of the current study.

Table 5

Means and Standard Error for Correct Completion of Math Problems in First and Last Condition

Treatment Condition	N	*M*	*SE*
First Condition	36	24.62*	2.80
Last Condition	36	35.06*	2.70

Note. * = significant difference between at the $p < .05$ level.

DISCUSSION

The current investigation evaluated the effectiveness of an academic intervention that involved both a reinforcement component (DRA) and a prompting component to increase compliance on an academic task. In addition, moderator variables were evaluated to understand how individual differences, such as preferences for either escape or attention, influenced intervention effectiveness. The results of this study replicated previous findings on behavioral interventions (Athens & Vollmer, 2010; Northup et al.,1997; Shirley et al., 1997; St. Peter-Pipkin et al., 2010; Vollmer et al., 1999) that have found these interventions to be robust, even when implemented with less than perfect integrity. This study also extends previous findings by evaluating how individual differences can influence treatment effectiveness when components of an intervention are omitted.

Treatment Effectiveness

Results from the current study revealed that results from all three intervention conditions (full treatment, reinforcement only, prompting only) were significantly different from the baseline condition, indicating that all three intervention conditions were effective in increasing correct completion of math problems over the baseline condition. In general, for these participants implementing only one component of the intervention was as effective as implementing the entire intervention package. In addition, there were no significant differences between any treatment condition, meaning that for these participants the reinforcement only and prompting only conditions were as effective as the full treatment condition. These findings are similar to the findings of Gansle and McMahon (1997), which found self-monitoring to be effective even when components, such as graphing and reinforcement, were not implemented.

Both studies provide some evidence that commonly used behavioral interventions may not need to be implemented with 100% integrity to change behaviors such on-task behavior or noncompliance. Because these interventions have been found to be effective even when components are omitted, it provides some evidence that common behavior interventions may be overly complex without the benefit of increased effectiveness.

These findings are of particular importance because behavioral interventions often involve multiple components, which can make them complex and more vulnerable to treatment integrity failures and more difficult to implement (Gresham, 1989). Therefore, understanding which components are necessary can provide valuable information in designing general behavioral interventions. For the participants included in this study the combined intervention was as effective as reinforcement only and prompting only. Because all of the conditions were equally effective individual components could be implemented or omitted based on individual preferences. For example, a teacher may choose to implement only the prompting component due to its ease and the fact that it requires few materials. On the other hand, a teacher may choose to implement the reinforcement component because reinforcement based procedures are often view as ethically superior to interventions without a reinforcement component.

Reinforcer Assessment

The current study replicated the use of a reinforcer assessment to determine preferences toward escape or attention (St. Peter-Pipkin et al, .2010) as well as extended the use of this tool to a larger participant group and a new population. There is abundant research demonstrating the use of function based interventions to both increase appropriate behavior

and decrease problem behavior (Carr & Durand, 1985; Derby et al., 2000; Hanley, Iwata, & McCord, 2003) with an experimental FA as gold standard for identifying function (O'Neill, Horner, Albin, Storey, & Sprague, 1996). Although experimental FA's have been empirically validated as an effective tool to identify function they are often not practical to implement in the natural environment (e.g., educational and home settings). Common concerns regarding implementing experimental FA's include the inability to properly control extraneous variables, the need to have highly trained personal to oversee and conduct the assessment (Tarbox et al., 2009), as well as the necessity to observe problem behavior. Alternative methods of identifying function have included the use of indirect assessments such as the Motivation Assessment Scale (MAS) by Durand and Crimmins (1987) or descriptive assessments such as A-B-C analysis. Although these alternative assessments use fewer resources, they have also been found to have inconsistent reliability (Iwata, Kahng, Wallace, & Lindberg, 2000) and to produce inaccurate results when compared to an experimental FA (Hall, 2005; Lerman and Iwata 1993). When considering the limitations of both experimental FA's and indirect/descriptive assessments, finding an alternative method for identifying function in more natural environments seems warranted.

The assessment used in this investigation may provide a quick alternative to indirect and descriptive assessments. It is possible that this reinforcer assessment may produce more reliable outcomes because it requires the individual to actively engage in choice between two potential reinforcers, although this is speculative at this point and warrants further investigation. That is, this reinforcer assessment would need to be compared to an

experimental functional analysis to validate it is as a reliable measure as well as replicate it's use with individuals who have a history of problem behavior.

Break Group. The findings from the current study indicated that participants who preferred escape responded differently than the participants in the ambiguous or attention groups in the assessed treatment conditions. The participants who completed more than 50% of correct problems on the "break" worksheet completed fewer problems in baseline (*M*=11.41) compared to the ambiguous group (*M*=18.91) and the attention group (*M*=28.92). For these participants the full treatment (*M*=32.08) was most effective in increasing correct completion of math problems followed by the reinforcement only condition (*M*=25.17) and prompting only condition (*M*=24.75) respectively. The one-way repeated measures ANOVA yielded a significant difference between the full treatment and reinforcement only condition for the break group; however the difference between the means of reinforcement only and prompting only conditions is less than 1 and the lack of significance found in the latter is likely due to the large amount of variability found in the data. This finding differs from the findings of the overall ANOVA, which found no differences between any of the treatment conditions.

Participants who preferred a break over attention benefited most from both reinforcement and prompting. However, there was no difference in responding when each component was implemented in isolation. This finding suggests a clear benefit to the combined intervention for these individuals. It is also interesting to note that when only one component was implemented, responding was nearly the same regardless of which component was included (i.e., reinforcement or prompting). This effect has also been demonstrated in previous studies that have shown a more positive outcome when multiple treatment components are

combined (Mazaleski et al., 1993; Rogers-Warren, Warren, & Baer, 1977; Eckert, Ardoin, Daly, & Martens,2002; Wacker, et al., 1990). Because this finding was only true for the individuals who preferred a break over attention, it highlights the necessity to further understand when and with whom intervention complexity yields a more positive outcome that outweighs any challenges associated with having multiple intervention components.

For the break group, the prompting component could be conceptualized as a function based intervention since it was the part of the intervention that directly targeted escape by not allowing the participant to escape the task (i.e., escape extinction). In contrast, the reinforcement only condition allowed for escape because there were no consequences for not working and provided reinforcement only for correct math problem completion. This finding is consistent with research that has demonstrated positive reinforcement to be effective for the treatment of problem behavior maintained by negative reinforcement. Lalli et al., (1999) demonstrated that positive reinforcement delivered contingently was effective at both decreasing problem behavior and increasing compliance. In addition Lomas, Fisher, and Kelley, (2010) demonstrated that positive reinforcement delivered noncontiengently can reduce problem behavior and increase compliance without the use of extinction. Lomas et al. demonstrated that this decrease in problem behavior and increase in compliance was likely due to the positive reinforcer acting as an abolishing operation (Laraway, Snycerski, Michael, & Poling, 2003). That is, providing access to positive reinforcement decreased participants' motivation to escape the task. It is possible that the same is true for the participants in the current study; access to the positive reinforcer (i.e., token) decreased their motivation to escape, and compliance therefore increased.

Ambiguous Group and Attention Group. The reinforcer assessment indentified 24 participants who allocated 50% or more of their responding to the "talk" worksheet, which represented two-thirds of the participants included in this study. These 24 participants were divided into two groups in order to analyze their data separately. Data analysis revealed that all treatment conditions were superior to the baseline condition for the ambiguous group, indicating that for these participants there was no advantage to the combination of reinforcement and prompting and components was as effective alone as they were in combination. The results for the attention group differed slightly in that for these participants there was not a statistically significant difference between any conditions, including the baseline condition. The participants included in the attention group worked regardless of the contingences in place and more than likely represented a group of children who would not need a behavioral intervention to address compliance.

Condition Carryover Effects

Repeated measures designs are particularly vulnerable to carryover effects because each participant is exposed to multiple conditions (Brooks, 2012); carryover occurs when exposure to one condition influences responding in subsequent conditions. The most common method to control for carryover effects is to counterbalance condition order across participants. The assumption is that counterbalancing will protect against the effects of exposure to a condition that subsequently affects responding in a later condition by distributing this effect across participants. Reese, 1997 argued that this assumption is not always correct and investigators should systematically evaluate carryover effects to insure correct interpretation of the data.

In the current study, carryover effects were evaluated to understand how exposure to reinforcement or prompting may have influenced responding in subsequent conditions. This evaluation was particularly important because condition order was counterbalanced across participants such that each condition was just as likely to be the second, third or fourth condition as it was to be the first (Reese, 1997). To understand how reinforcement affected baseline means, means of participants who completed the baseline condition before any reinforcement condition were compared to means of participants who completed the baseline condition following a reinforcement condition. The results yielded a significant difference between the groups, with a higher mean of correct problem completion from participants who experienced reinforcement prior to baseline than participants who experienced baseline prior to reinforcement. This same result occurred when prompting was evaluated. That is, the mean of participants who experienced prompting prior to baseline was significantly higher than the mean of participants who experienced baseline prior to a prompting. These results provide some evidence that condition order may have influenced outcomes in subsequent conditions, and more specifically exposure to treatment, regardless of type, influenced responding in the baseline condition.

Additional evidence of carryover effects was provided by analyzing means of the first and last condition. Participants completed significantly more correct problems in the last condition than in the first. It is possible this may have been due to exposure to treatment in previous conditions because in all cases the participants would have experienced at least two treatment conditions before the last condition. It is impossible to understand how each condition influenced responding in a subsequent condition, but these analyses provide

evidence that order and experiment design may have impacted the results of this study. Because of these carryover effects results should be interpreted with caution.

A second way in which condition order was evaluated was by comparing the first exposure to a treatment component to the second exposure to the same treatment component. On average participants completed fewer problems correctly the first time they were exposed to reinforcement and completed more problems correctly the second time they were exposed to reinforcement. This difference was found to be significant, indicating that prior exposure to reinforcement increased reinforcement effectiveness. On the other hand, this effect was not found for prompting. This effect with reinforcement may have confounded the results of this study; that is exposure to reinforcement in previous conditions may have influenced responding on future conditions. Although this effect may have influenced the results of the current investigation this finding provides some evidence that there may be some clinical advantages to reinforcement over prompting because the effects of reinforcement appear to be more robust than the effects of prompting.

Reinforcement history effect, sometimes referred to as a carryover effect, is when a behavior pattern persists that is similar to when reinforcement was in effect once reinforcement has been terminated. This effect has been well documented in the basic research (Freeman & Lattal, 1992; Ono & Iwabuchi, 1997; Urbain, Poling, Millam, & Thompson, 1978) and has been replicated in the applied literature (Mace, Neef, Shade, & Mauro, 1994; Martens, Bradley, & Eckert, 1997; Martens, Hilt, Needham, Sutterer, Panahon & Lannie, 2003;Weiner, 1964). In the study conducted by Martens et al. (1997), the authors evaluated carryover effects of both reinforcement and prompting by exposing participants to three

different reinforcement histories that included either reinforcement or a combination of prompts and reinforcement. Participants were exposed to these reinforcement contingencies for only two minutes followed by 8 min of extinction. Both participants in this study continued to complete school worksheets during extinction after exposure to a reinforcement condition, with the highest level of responding following the condition that included both reinforcement and prompting. It is possible that similar carryover effects were observed in the current study; that is, even though participants were exposed to only a brief condition with reinforcement and/or prompting the increased responding observed in later conditions may have been carryover from this exposure to reinforcement and/or prompting.

Limitations and Future Directions

Participants for this study were recruited from regular education classes and may or may not have had a history of noncompliance in academic settings. This fact is important for the generalizability of these finding to high risk populations with a history of noncompliance. Therefore, future investigations should evaluate similar interventions with at-risk populations or with individuals that are referred for non-compliance. Another limitation is the lack of additional information collected on each of the participants. Additional information such as current work habits in the classroom, history of noncompliance and/or grades in math could have been used to validate the finding of the functional reinforcer assessment.

Carryover effects were a significant limitation in the current study. It is clear from the follow-up t-test that exposure to one condition affected responding in subsequent conditions. Future studies should consider alternative methods for reducing carryover. In the current study, participants were exposed to all four conditions within the same session. Future studies could

conduct only one condition per session or per day to increase the amount of time that elapses between conditions. Future studies could also consider conducting the baseline condition first, then counterbalance remaining treatment conditions to eliminate reinforcement history from affecting responding in the baseline condition. An across participants rather than within participants design could be used to eliminate carryover effects. However, this alternative design would require a significant increase in the number of participants as well as prevent evaluation of within participant differences to each treatment condition.

The current study was also conducted in an analog setting and not in the classroom. It also used a contrived a situation to increase the likelihood of noncompliance (i.e., by allowing free access to a movie throughout each conditions). Although this method was successful in that most participants were more likely to watch a movie rather than complete math worksheets when reinforcement and/or prompting was not provided, it may not be representative of what would occur in the participants' natural environment. Therefore, replication in a classroom is needed to evaluate treatment effectiveness and assess generalizability of the current study's findings.

Another limitation of the current study was the number of participants and the distribution of responses in the reinforcer assessment. The current study only included 36 participants and therefore should be replicated with a larger number of participants. A larger participant pool would also provide more power for group data analysis. In the current study, the 36 participants were divided into 3 groups of 12. Having larger number of participants in each group may reveal additional differences between groups.

Another limitation of the current study is that two-thirds of the participants were identified as having a preference for attention over a break. Had the results been more evenly distributed it would have provided a more powerful analysis of group results. In addition, the break, ambiguous and attention groups were somewhat arbitrarily created. The ambiguous group, which consisted of participants who choose attention over a break 50%-89% may not represent a true ambiguous group. Having more evenly distributed groups would allow for the creation of three distinct groups; a group who prefers breaks over attention, a group with no preference for either or an equal preference for both, and a group who prefers attention over break. Therefore future studies should replicate these methods but increase in overall number of participants.

Future investigations should compare the results of reinforcer assessments with results from an experimental FA to evaluate its reliability and validity with more standard measures. This step is necessary to establish this tool as a reliable method for indentifying potential functional reinforcers. If this tool reliably predicts which reinforcer maintains problem behavior it could be used in variety of settings in which experimental FAs are more difficult to implement, such as schools and homes. In addition, because this tool does not require the individual to engage in problem behavior it could be used with individuals who engaged in less frequent problem behavior, with individuals who engage in life threatening behavior, or with individuals who engaged in covert behaviors. This assessment should also be replicated with individuals who are both typically developing as well as individuals with developmental disabilities to validate this measure across a variety of populations.

The current study found that a common behavior intervention that included both a reinforcement and prompting component was effective in increasing correct completion of math problems and the intervention remained effective even when reinforcement or prompting failed to be implemented. In addition, a reinforcer assessment successfully indentified participants with preferences for escape or attention. Participants in the break group benefited most from the full intervention, which included reinforcement and prompting. On the other hand there was no added benefit of the full intervention, with all treatment conditions equally effective, for participants who were in the ambiguous group. For the participants in the attention group there were no specific benefits to any treatment condition; these participants were observed to work regardless of the treatment components in place. This finding replicates previous findings that DRAs may not need to implement with perfect integrity to remain effective (Northup et al., 1997; Vollmer et al., 1999; St. Peter-Pipkin et al., 2010). In addition, the current study extended findings by using a reinforcer assessment to evaluate preferences for either escape or attention and how those preferences may impact treatment effectiveness. These findings taken together highlight the need for more research to evaluate treatment components necessary for treatment effectiveness in conjunction with individual differences which may also impact treatment effectiveness.

REFERENCES

Athens, E. S., & Vollmer, T.R. (2010). An investigation of differential reinforcement of alternative behavior without extinction. *Journal of Applied Behavior Analysis, 43,* 569-589.

Austin, J.L., & Soeda, J.M. (2008). Fixed-time teacher attention to decrease off-task behaviors of typically developing third graders. *Journal of Applied Behavior Analysis, 41,* 279-283.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97.

Brooks, J.L. (2012). Counterbalancing for serial order carryover effects in experimental condition order. *Psychological Methods, 17,*600-614.

Brosnan, J., & Healy, O. (2011). A review of behavioral interventions for the treatment of aggression in individuals with developmental disabilities*. Research in Developmental Disabilities: A Multidisciplinary Journal, 32*, 437-446.

Bryan, L.C., & Gast, D.L. (2000). Teaching on-task and on-schedule behaviors to high-functioning children with autism via picture activity schedules. *Journal of Autism and Developmental Disorders, 30*, 553-567.

Campbell, A., & Anderson, C. M. (2011). Check-in/check-out: A systematic evaluation and component analysis. *Journal of Applied Behavior Analysis, 44,* 315-326.

Carbone, V.J., Sweeney-Kerwin, E.J., Attanasio, V., Kasper, T. (2010). Increasing the vocal responses of children with autism and developmental disabilities using manual sign mand training and prompt delay. *Journal of Applied Behavior Analysis, 43,* 705-709.

Carr, E.G., & Durand, V.M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18,* 111-126.

Cataldo, M. F., Bessman, C. A., Parker, L. H., Pearson, J. E. R., & Rogers, M. C. (1979). Behavioral assessment for pediatric intensive care units. *Journal of Applied Behavior Analysis, 12*, 83-97.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cooper, J.O., Heron, T.E., & Heward, W.L. (2007). *Applied behavior analysis* (2nd ed). Upper Saddle River, NJ: Prentice Hall.

Deitz, D.E.D., & Repp, A.C. (1983). Reducing behavior through reinforcement. *Exceptional Education Quarterly, 3,* 33-46.

Derby, K. M., Hagopian, L., Fisher, W. W., Richman, D., Augustine, M., Fahs, A., & Thompson, R. (2000). Functional analysis of aberrant behavior through measurement of separate response topographies. *Journal of Applied Behavior Analysis, 33*, 113-117.

Derby, K. M., Wacker, D. P., Berg, W., DeRaad, A., Ulrich, S., Asmus, J., Harding, J., Prouty, A., Laffey, P., & Stoner, E. A. (1997). The long-term effects of functional communication training in home settings. *Journal of Applied Behavior Analysis, 30*, 507-531.

DuPaul**,** G. J., Ervin, R. A., Hook, C. L., & McGoey, K. E.(1998). *Journal of Applied Behavior Analysis ,31*, 579-592.

Durand, V. M., & Crimmins, D. B. (1988). Identifying variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18*, 99–117.

Eckert, T. L., Ardoin, S.P., Daly, E. J., & Martens, B.K. (2002). Improving oral reading fluency: A brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Autism and Developmental Disorders, 35*, 271–281.

Falcomata, T.S., Ringdahl, J.E., Christensen, T.J., & Boelter, E.W. (2010). An evaluation of prompt schedules and mand preference during functional communication training. *The Behavior Analyst Today, 11,* 77-84.

Fischer, S. M., Iwata, B. A., & Mazaleski, J. L. (1997). Noncontingent delivery of arbitrary reinforcers as treatment for self-injurious behavior. *Journal of Applied Behavior Analysis, 30*, 239-249.

Fisher, W., Piazza, C., Cataldo, M., Harrell, R., Jefferson, G., & Conner, R. (1993). Functional communication training with and without extinction and punishment. *Journal of Applied Behavior Analysis, 26*, 23-36.

Flood, W. A., Wilder, D. A., Flood, A. L., & Masuda, A. (2002). Peer-mediated reinforcement plus prompting as treatment for off-task behavior in children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis, 35*, 199-204.

Freeland, J.T., & Noell, G.H. (1999). Maintaining accurate math responses in elementary school students: the effects of delayed intermittent reinforcement and programming common stimuli. *Journal of Applied Behavior Analysis*, *32*, 211-215.

Freeman, T.J., & Lattal, K.A. (1992). Stimulus control of behavioral history. *Journal of the Experimental Analysis of Behavior, 57*, 5-15.

Gansle, K.A., & McMahon, C.M. (1997). Component integrity of teacher intervention management behavior using a student self-monitoring treatment: An experimental analysis. *Journal of Behavioral Education, 7*, 405-419.

Gardner, A. W., Wacker, D.P., & Boelter, E.W. (2009). An evaluation of the interaction between quality of attention and negative reinforcement with children who display escape-maintained problem behavior. *Journal of Applied Behavior Analysis, 42,* 343-348.

Greene, B. F., Bailey, J. S., & Barber, F. (1981). An analysis and reduction of disruptive behavior on school buses. *Journal of Applied Behavior Analysis, 14*, 177-192.

Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review, 18*, 37-50.

Gresham, F.M. (2011). Social Behavioral Assessment and Intervention: Observations and Impressions. *School Psychology Review, 40*, 275-283.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis, 26*, 257–263.

Gresham, F.M., Gansle, K.A., Noell, G.H., Cohen, S. & Rosenblum, S. (1993). Treatment integrity of school-based behavioral intervention studies: 1980-1990. *School Psychology Review, 22,* 254-272.

Hagopian, L.P., Kuhn, D.E., &. Strother, G.E. (2009). Targeting social skills deficits in an adolescent with pervasive developmental disorder. *Journal of Applied Behavior Analysis, 42,* 907-911.

Hagopian, L. P., Fisher, W. W., & Legacy, S. M. (1994). Schedule effects of noncontingent reinforcement on attention-maintained destructive behavior in identical quadruplets. *Journal of Applied Behavior Analysis, 27*, 317-325.

Hall, S. S. (2005). Comparing descriptive, experimental, and informant-based assessments of problem behaviors. *Research in Developmental Disabilities, 26*, 514–526.

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147-185.

Holcombe, D.M., Wolery, M., & Snyder, E. (1994). Effects of two levels of procedural fidelity with constant time delay on children's learning. *Journal of Behavioral Education, 4*, 49-73.

Ingham, R. J., & Andrews, G. (1973). An analysis of a token economy in stuttering therapy. *Journal of Applied Behavior Analysis, 6*, 219-229.

Iwata, B.A., Dorsey, M.F., Slifer, K.J., Bauman, K.E., & Richman, G.S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197-209. (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2,* 3-20, 1982)

Iwata, B. A., Kahng, S., Wallace, M. D., & Lindberg, J. S. (2000). The functional analysis model of behavior assessment. In J. Austin & J. E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 61–89). Reno: Context.

Iwata, B.A., Pace, G.M., Cowdery, G.E., & Mitlenberger, R.G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis, 27*, 131-144.

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R.G., …Willis, K.D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *Journal of Applied Behavior Analysis*, *27*, 215-240.

Johnston, J., & Pennypacker, H., (1980). *Strategies and tactics in human behavioral research.* Hillsdale, NJ: Erlbaum.

Jones, K. M., Drew, H. A., & Weber, N. L. (2000). Noncontingent peer attention as treatment for disruptive classroom behavior. *Journal of Applied Behavior Analysis, 33*, 343-346.

Kahng, S., Iwata, B. A., DeLeon, I. G., & Worsdell, A. S. (1997). Evaluation of the "control over reinforcement" component in functional communication training. *Journal of Applied Behavior Analysis, 30*, 267-277.

Karsten, A.M., & Carr, J.E. (2009).The effects of differential reinforcement of unprompted responding on the skill acquisition of children with autism. *Journal of Applied Behavior Analysis, 42*, 327-334.

Kennedy, C. H., Meyer, K. A., Knowles, T., & Shukla, S. (2000). Analyzing the multiple functions of stereotypical behavior for students with autism: Implications for assessment and treatment. *Journal of Applied Behavior Analysis, 33*, 559-571.

Kleinman, K. E., & Saigh, P. A. (2011). The Effects of the Good Behavior Game on the Conduct of Regular Education New York City High School Students. *Behavior Modification, 35*, 95-105.

Kuhn, D.E., DeLeon, I.G., Fisher, W.W., & Wilke, A.E. (1999). Clarifying an ambiguous functional analysis with matched and mismatched extinction procedures. *Journal of Applied Behavior Analysis, 32*, 99-102.

Lalli, J. S., & Casey, S. D. (1996). Treatment of multiply controlled problem behavior. *Journal of Applied Behavior Analysis, 29,* 391-396.

Lalli, J. S., Casey, S., Goh, H., & Merlino, J. (1994). Treatment of escape-maintained aberrant behavior with escape extinction and predictable routines. *Journal of Applied Behavior Analysis, 27*, 705-714.

Lalli, J.S., Vollmer, T. R., Porgar, P. R., Wright, C., Borrero, J., Daniel, D., …May. J.(1999). Competition between positive and negative reinforcement in the treatment of escape behavior. *Journal of Applied Behavior Analysis*, *32*, 285-296.

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*, 407-414.

Lerman, D. C.,& Iwata, B. A.(1993). Descriptive and experimental analysis of variables maintaining selfinjurious behavior. *Journal of Applied Behavior Analysis, 26*, 293–319.

Lerman, D. C.,& Iwata, B. A.(1995). Prevalence of the extinction burst and its attenuation during treatment. *Journal of Applied Behavior Analysis, 28*, 93-94.

Lomas, J. E., Fisher, W.W., & Kelley, M.E. (2010). The Effects of Variable-time Delivery of Food Items and Praise on Problem Behavior Reinforced by Escape. *Journal of Applied Behavior Analysis, 43,* 425-435.

Mace, F.C., Neef, N.A., Shade, D., & Mauro, B.C. (1994).Limited matching on concurrent-schedule reinforcement of academic behavior. *Journal of Applied Behavior Analysis, 27,* 585-596.

Marcus, B. A., & Vollmer, T. R. (1995). Effects of differential negative reinforcement on disruption and compliance. *Journal of Applied Behavior Analysis, 28*, 229-230.

Martens, B.K., Bradley, T.A., & Eckert, T.L. (1997). Effects of reinforcement history and instructions on the persistence of student engagement. *Journal of Applied Behavior Analysis, 30*, 569-572.

Martens, B.K., Hilt, A.M., Needham, L.R., Sutterer, J.R., Panahon, C.J., & Lannie, A.L. (2003). Carryover effects of free reinforcement on children's work completion. *Behavior Modification, 27,*560-577.

Mayfield, K. H., & Vollmer, T. R. (2007). Teaching math skills to at-risk students using home-based peer tutoring. *Journal of Applied Behavior Analysis, 40*, 223-237.

Mazaleski, J.L., Iwata, B.A., Vollmer, T.R., Zarcone, J.R., & Smith, R.G. (1993). Analysis of the reinforcement and extinction components in DRO contingencies with self-injury. *Journal of Applied Behavior Analysis, 26*, 143-156.

McIntyre, L., Gresham, F., DiGennaro, F., & Reed, D. (2007). Treatment integrity of school-based interventions with children in the journal of applied behavioral analysis 1991–2005. *Journal of Applied Behavioral Analysis,40*, 659–672.

National Reading Panel (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implication for reading instruction.* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Noell**,** G. H., Connell, J. E., & Duhon, G. J. (2006). Spontaneous response generalization during whole word instruction: reading to spell and spelling to read. *Journal of Behavioral Education, 15*, 121-130.

Noell, G.H., Gresham, F.M., & Gransle, K.A. (2002). Does treatment integrity matter: A preliminary investigation of instruction implementation and mathematics instruction. *Journal of Behavioral Education, 11*, 51-67.

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly, 12,* 77-88.

Noell, G.H., Witt, J.C., Slider, N.J., Connell, J.E., Gatti, S.L., Williams, K.L., …Duhon, G. J. (2005). Treatment implementation following behavioral consolation in schools. A comparison of three follow-up strategies. *School Psychology Review, 34,* 87-106.

Northup, J., Fisher, W., Kahng, S., Harrel, B., & Kurtz. P. (1997). An assessment of the necessary strength of behavioral treatments for severe behavior problems. *Journal of Behavioral Education, 9*, 1-16.

O'Brien, T. P., Riner, L. S., & Budd, K. S. (1983). The effects of a child's self-evaluation program on compliance with parental instructions in the home. *Journal of Applied Behavior Analysis, 16,* 69-79.

O'Neill, R. E., Horner, R. H., Albin, R. W., Storey, K., & Sprague, J. R. (1996). Functional assessment and program development for problem behavior: A practical handbook. Wadsworth.

Ono, K., & Iwabuchi, K. (1997). Effects of histories of differential reinforcement of response rate on variable-interval responding. *Journal of the Experimental Analysis of Behavior, 67*, 311-322.

Peterson, L., Homer, A., & Wonderlich, S. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis, 15,* 477-492.

Piazza, C. C., Bowman, L. G., Contrucci, S. A., Delia, M. D., Adelinis, J. D., & Goh, H. (1999). An evaluation of the properties of attention as reinforcement for destructive and appropriate behavior. *Journal of Applied Behavior Analysis, 32*, 437-449.

Piazza, C. C., Moes, D. R., & Fisher, W. W. (1996). Differential reinforcement of alternative behavior and demand fading in the treatment of escape-maintained destructive behavior. *Journal of Applied Behavior Analysis, 29*, 569-572.

Poling, A., & Normand, M. (1999). Noncontingent reinforcement: An inappropriate description of time-based schedules that reduce behavior. *Journal of Applied Behavior Analysis, 32*, 237-238.

Reese, H.W. (1997). Counterbalancing and other uses of repeated-measures latin-square designs: Analyses and interpretations. *Journal of Experimental Child Psychology, 64*,197-158.

Roberts, R. N., Nelson, R. O., & Olson, T. W. (1987). Self-instruction: An analysis of the differential effects of instruction and reinforcement. *Journal of Applied Behavior Analysis, 20*, 235-242.

Roane, H. S., Fisher, W. W., & Sgro, G. M. (2001). Effects of a fixed-time schedule on aberrant and adaptive behavior. *Journal of Applied Behavior Analysis, 34*, 333-336.

Roane, H. S., Fisher, W. W., Sgro, G. M., Falcomata, T. S., & Pabico, R. R. (2004). An alternative method of thinning reinforcer delivery during differential reinforcement. *Journal of Applied Behavior Analysis, 37,* 213-218.

Rogers-Warren, A., Warren, S. F., & Baer, D. M. (1977). A component analysis: Modeling, self-reporting, and reinforcement of self-reporting in the development of sharing. *Behavior Modification, 1*, 307-322.

Salend, S. J. (1983). Minimizing integrity-of-treatment violations. *Exceptional Children, 50*, 179-181.

Sanetti, L. M. H., Gritter, K. L., & Dobey, L. (2011). Treatment integrity of interventions with children in the school psychology literature from 1995 to 2008. *School Psychology Review, 40,* 72-84.

Sasso, G. M., Reimers, T. M., Cooper, L. J., Wacker, D., Berg, W., Steege, M., … Allaire, A. (1992). Use of descriptive and experimental analysis to identify the functional properties of aberrant behavior in school settings. *Journal of Applied Behavior Analysis, 25*, 809-821.

Schieltz, K. M., Wacker, D. P., Harding, J.W., Berg, W. K., Lee, J. F., Padilla, D., ...Ibrahimovic, M. (2011). Indirect effects of functional communication training on non-targeted disruptive behavior. *Journal of Behavioral Education, 20,* 15-32.

Stahr, B., Cushing, D., & Lane, E. (2006). Efficacy of a function-based intervention in decreasing off-task behavior exhibited by a student with ADHD. *Journal of Positive Behavior Interventions*, *8*, 201-211.

Shirley, M.J., Iwata, B.A., Kahng, S., Mazaleski, J.L., & Lerman, D.C. (1997). Does functional communication training compete with ongoing contingencies of reinforcement? An analysis during response acquisition and maintenance. *Journal of Applied Behavior Analysis, 30*, 93-104.

St.Peter Pipkin, C., Vollmer, T.R., & Sloman, K.N. (2010). Effects of treatment integrity failures during differential reinforcement of alternative behavior: A translational model. *Journal of Applied Behavior Analysis, 43*, 47-70.

Tarbox, J., Wilke, A.E., Najdowski, A.C., Findel-Pyles, R.S., Balasanyan, S., Caveney, A.C., Chilingaryan, V., …Tia, B. (2009). Comparing indirect, descriptive and experimental functional assessments of challenging behavior in children with autism. *Journal of Developmental and Physical Disabilities, 21,*493-514.

Thompson, R. H., Fisher, W. W., Piazza, C. C., & Kuhn, D. E. (1998). The evaluation and treatment of aggression maintained by attention and automatic reinforcement. *Journal of Applied Behavior Analysis, 31*, 103-116.

Touchette, P.E. (1971). Transfer of stimulus control: Measuring the moment of transfer. *Journal of the Experimental Analysis of Behavior, 15*, 347-354.

Urbain, C., Poling, A., Milliam, J., & Thompson, T. (1978). D-Amphetamine and fixed-interval performance: Effects of operant history. *Journal of the Experimental Analysis of Behavior, 29*, 385-392.

Vollmer, T.R., & Iwata, B.A. (1992). Differential reinforcement as a treatment for behavior disorders: Procedural and functional variations. *Research in Developmental Disabilities, 13,* 393-417.

Vollmer, T.R., Iwata, B.A., Zarcone, J.R., Smith, R.G., & Mazalski, J.L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis, 26,* 9-21.

Vollmer, T. R., Marcus, B. A., & Ringdahl, J. E. (1995). Noncontingent escape as treatment for self-injurious behavior maintained by negative reinforcement. *Journal of Applied Behavior Analysis, 28,* 15-26.

Vollmer, T.R., Roane, H.S., Ringdahl, J.E., & Marcus, B.A. (1999). Evaluating treatment challenges with differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis, 32,* 9-23.

Wacker, D. P., Steege, M. W., Northup, J., Sasso, G. M. ,Berg, W., Reimers, T., … Donn, L. (1990). A component analysis of functional communication training across three topographies of severe behavior problems. *Journal of Applied Behavior Analysis, 23,* 417-429.

Wahler, R.G., Vigilante, V.A., & Strand, P.S. (2004). Generalization in a child's oppositional behavior across home and school settings. *Journal of Applied Behavior Analysis, 37*, 43-51.

Walker, B. (2008). Constant and progressive time delay procedures for teaching children with autism: A literature review. *Journal of Autism and Developmental Disorders, 38,* 261-275.

Watson, T.S, Sterling H.M, & McDade A. (1997). Demystifying behavioral consultation. *School Psychology Review**,26,*** 467–474.

Weiner, H. (1964). Conditioning history and human fixed-interval performance. *Journal of the Experimental Analysis of Behavior, 7*, 383-385.

Wheeler, J., Baggett, B., Fox, J., & Blevins, L. (2006). Treatment integrity: A review of intervention studies conducted with children with autism. *Focus on Autism and Other Developmental Disabilities, 21*, 45–54.

Witt, J. C, Noell, G. H, LaFleur, L. H, & Mortenson, B. P.(1997). Teacher use of interventions in general education settings: Measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis*, *30*, 693–696.

Wood, B.K, Blair, K.C., & Ferro, J.B. (2009). Young Children with Challenging Behavior: Function-Based Assessment and Intervention. *Topics in Early Childhood Special Education, 29*, 68-78.

Worsdell, A. S., Iwata, B. A., Hanley, G. P., Thompson, R. H., & Kahng, S. (2000). Effects of continuous and intermittent reinforcement for problem behavior during functional communication training. *Journal of Applied Behavior Analysis, 33*, 167-179.

Yeaton, W.H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments; Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49,* 156-167.

Yilmaz, I., Konukman, F., Birkan, B., & Yanardag, M. (2010). Effects of most to least prompting on teaching simple progression swimming skill for children with autism. *Education and Training in Autism and Developmental Disabilities, 45(3),* 440-448.

APPENDIX
INSTITUTIONAL REVIEW BOARD APPROVAL

Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.8782
irb@lsu.edu | lsu.edu/irb

ACTION ON PROTOCOL APPROVAL REQUEST

TO: George Noell
Psychology

FROM: Robert C. Mathews
Chair, Institutional Review Board

DATE: March 6, 2012
RE: **IRB#** 3262

TITLE: Treatment Integrity Failures Matched to Behavioral Function

New Protocol/Modification/Continuation: New Protocol

Review type: Full ___ Expedited _X_ . **Review date:** 3/7/2012

Risk Factor: Minimal __X__ Uncertain _____ Greater Than Minimal ___

Approved_____X_____ **Disapproved**_____

Approval Date: 3/7/2012 . **Approval Expiration Date:** 3/6/2013

Re-review frequency: (annual unless otherwise stated)

Number of subjects approved: 30

Protocol Matches Scope of Work in Grant proposal: (if applicable)_____

By: Robert C. Mathews, Chairman Robert C. Mathews

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. SPECIAL NOTE:

**All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb*

VITA

Joanna E. Lomas Mevers is a candidate for the Doctor in Philosophy degree in the school psychology program at Louisiana State University. She graduated with her Bachelor of Science degree in psychology in 2001 from Louisiana State University and received her Masters of Arts degree in 2008 from Georgia State University. Joanna is currently completing her pre-doctoral internship at The Marcus Autism Center in Atlanta, Georgia. Joanna completed her graduate work under the supervision of Dr. George H. Noell.