DIFFERENTIAL REINFORCEMENT OF OTHER BEHAVIOR (DRO)
APPLIED CLASSWIDE WITH YOUNG CHILDREN

A dissertation submitted to the
Kent State University College and Graduate School
of Education, Health, and Human Services
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

by

Rosemarie Daddario

August 2010
© Copyright by Rosemarie Daddario 2010

All Rights Reserved
A dissertation written by

Rosemarie Daddario

B. S., The Ohio State University, 1998

M. Ed., Kent State University, 2005

Ph. D., Kent State University, 2010

Approved by

__________________________________, Director, Doctoral Dissertation Committee
Richard J. Cowan

__________________________________, Member, Doctoral Dissertation Committee
Karla Anhalt

__________________________________, Member, Doctoral Dissertation Committee
Lyle Barton

Accepted by

__________________________________, Interim Director, School of Lifespan
Mary Dellmann-Jenkins Development and Educational Sciences

__________________________________, Dean, College and Graduate School of
Daniel F. Mahony Education, Health, and Human Sciences
The purpose of this study was to further expand the evidence-based literature on the effectiveness of Differential Reinforcement of Other Behavior (DRO) as an isolated intervention when reducing challenging behaviors. This research study addressed the issues of reinforcement, determining interval length, and momentary versus whole DRO application to further explain these variables as indicated in previous research.

Specifically, this study conducted a reinforcement pre-assessment and calculated interval length (interresponse time) on a weekly basis. It examined the effectiveness of implementing whole interval DRO (wDRO) during treatment and momentary interval DRO (mDRO) during maintenance at the classwide level to reduce talking out in typically developing kindergarten children using a multiple baseline across settings design. Measures of treatment acceptability, treatment integrity, and other social validity measures were also utilized.

Results indicated that the intervention effectively reduced talking out in both settings, with average response rates during floor reduced from 4.3 responses per minute (rpm) to 3.2 rpm, and 3.9 rpm to 2.2 rpm during math. Although treatment acceptability scores were below the acceptable range, treatment integrity data revealed high adherence
to the DRO protocol. Social validity, partly measured by the behavior rating scales, indicated no statistical difference in the teacher’s or parents’ reports of behavior intensity. Parents reported a statistical difference in the extent they saw the behavior as problematic at home.

Despite some observed limitations, several implications appear warranted. First, the relative success of momentary interval DRO over whole interval DRO appears to contradict past research. Second, it is important for practitioners to note the crucial issue of teacher acceptability and feasibility. Finally, the value of varying powerful reinforcement and delivering that reinforcement immediately after the behavior continues to demand attention. Implications for research and practice are included.
ACKNOWLEDGMENTS

I would first like to thank my husband, Shane, for putting up with the emotional ups and downs of applied research, for helping me brainstorm ideas on how to improve the intervention, for listening to me drone on about data, for listening to me practice my defense presentations, and most importantly, for telling me that you’re proud of me. I would like to thank my little girls, Ella and Ava, for letting Mommy do her schoolwork. I know you don’t understand the importance now, but one day I will enjoy explaining to you why Mommy is called a doctor, but she can’t fix a broken arm like Dr. George can. Thank you to my parents, who laid the foundation for me to believe that I can accomplish anything and who encouraged me every step of the way. Thank you for all of the babysitting and dinners over the past few years to make this dream possible.

I would like to thank my data researchers, Emilee and Brooke, who took away the stress of worrying about reliable, responsible data collection. Thank you for a job well done! Thank you to my colleagues and friends, Julie and Brooke, for finding me participants, for helping me brainstorm ideas, for your thoughtful input, patience with my questions, and for keeping me focused. Thank you to Dawn for volunteering herself and her classroom for this study and for “hanging in there” until the end.

Lastly, but most importantly, I would like to thank my committee, Dr. Richard Cowan, Dr. Karla Anhalt, and Dr. Lyle Barton for your unending time, materials, guidance, thoughtful feedback, encouragement, and for providing a calm voice during my intervention meltdowns. Thank you for helping make this dream come true.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>I. LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>Applied Behavior Analysis</td>
<td>1</td>
</tr>
<tr>
<td>Antecedent &amp; Consequential Considerations</td>
<td>3</td>
</tr>
<tr>
<td>Antecedent Considerations</td>
<td>3</td>
</tr>
<tr>
<td>Consequential Considerations</td>
<td>4</td>
</tr>
<tr>
<td>Functional Assessment</td>
<td>4</td>
</tr>
<tr>
<td>History of ABA in Education</td>
<td>6</td>
</tr>
<tr>
<td>Positive Behavior Supports</td>
<td>7</td>
</tr>
<tr>
<td>Empirical Support</td>
<td>8</td>
</tr>
<tr>
<td>Links to Applied Behavior Analysis</td>
<td>10</td>
</tr>
<tr>
<td>Continuum of Support</td>
<td>10</td>
</tr>
<tr>
<td>Overview of Antecedent Strategies</td>
<td>12</td>
</tr>
<tr>
<td>Overview of Consequential Strategies</td>
<td>12</td>
</tr>
<tr>
<td>Punishment Procedures</td>
<td>13</td>
</tr>
</tbody>
</table>
Data Collection ........................................................................................................ 46

Single Subject Research Design .............................................................................. 49
  Social Validity/Social Comparison ........................................................................ 50
  Baseline ................................................................................................................ 50
  Pre-Intervention ...................................................................................................... 51
  Intervention/Whole DRO ...................................................................................... 52
  Pre-Maintenance Meeting .................................................................................... 53
  Maintenance/Momentary DRO/Reversal .............................................................. 54
  Post-Maintenance Meeting .................................................................................. 54

Measures ..................................................................................................................... 55
  Social Validity/Treatment Acceptability ............................................................... 55
  Treatment Integrity ............................................................................................... 56
  Social Validity of Goals ......................................................................................... 57
  Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R) ......................... 58
  Eyberg Child Behavior Inventory (ECBI) ............................................................ 58

Data Analysis ............................................................................................................. 59

III. RESULTS ................................................................................................................. 63
  Visual Analysis ........................................................................................................ 63
  Setting One ............................................................................................................. 64
  Setting Two ............................................................................................................ 66
  Across Settings ...................................................................................................... 66
Treatment Integrity .................................................................................................. 67
Social Validity Measures .......................................................................................... 68
Acceptability ........................................................................................................... 68
Goals ......................................................................................................................... 69
Comparison .............................................................................................................. 69
Supplementary Data ................................................................................................. 69
Interobserver Agreement ......................................................................................... 69
SESBI-R .................................................................................................................... 72
ECBI .......................................................................................................................... 72
Intervals of Time ....................................................................................................... 73
Conclusion .................................................................................................................. 74

IV. DISCUSSION .......................................................................................................... 75
Findings ...................................................................................................................... 75
Educational Links ....................................................................................................... 76
Hypotheses .................................................................................................................. 77
Teacher Acceptability ............................................................................................... 77
Treatment Integrity ................................................................................................... 80
Relationship between Treatment Acceptability, Integrity, & Outcomes ........... 81
SESBI-R/ECBI Implications ....................................................................................... 81
Participants ................................................................................................................ 82
Reinforcement ............................................................................................................. 83
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Antecedent Interventions</td>
<td>13</td>
</tr>
<tr>
<td>2. PND Intervention Effectiveness Ratings</td>
<td>61</td>
</tr>
<tr>
<td>3. Comparison of Mean Response Rate of Talking Out</td>
<td>70</td>
</tr>
<tr>
<td>4. IOA Results for the Target Class</td>
<td>71</td>
</tr>
<tr>
<td>5. IOA Results for the Social Comparison Class</td>
<td>71</td>
</tr>
<tr>
<td>6. Descriptive Data for SESBI-R and ECBI Behavior Rating Scales</td>
<td>73</td>
</tr>
<tr>
<td>7. Interresponse Times (IRT) for Floor</td>
<td>74</td>
</tr>
<tr>
<td>8. Teacher Variables That Influence Treatment Acceptability</td>
<td>78</td>
</tr>
</tbody>
</table>
CHAPTER I
LITERATURE REVIEW

Applied Behavior Analysis

The methodology and experimental approach of Applied Behavior Analysis (ABA) evolved from the experimental analysis of behavior introduced by Skinner (Cooper, Heron, & Heward, 2007). Skinner’s work concluded that behaviors were both influenced by antecedents (respondent) and consequences (operant). This discovery established the new science of experimental analysis of behavior. Skinner originally used rats and pigeons in laboratories as subjects; however, one of the first human studies was reported as early as 1949 by P. R. Fuller. At this point, researchers’ main goals were to determine whether the basic principles of behavior that were discovered with nonhumans also applied to humans. This laid the foundation for ABA in that later research began to use the principles of behavior to improve socially important behaviors in people (Cooper et al.).

Applied Behavior Analysis (ABA) is defined as the systematic or scientific application of behavioral or operant psychology to solve problems of social importance (Baer, Wolf, & Risley, 1968). It is an approach to change behavior using scientifically based procedures that are based on principles of learning (Kearney, 2008). Those implementing ABA are following a ‘scientific method’ in that conditions are controlled in order to identify a relationship. It is this process that attempts to establish a functional
relationship between a behavior and its consequence and therefore, evaluate the success of an intervention (Alberto & Troutman, 2009).

The process is considered *applied* because it is applied to affect behavior that is of interest to society and of importance to those involved (Baer et al., 1968). The behavior being manipulated must be quantified, measured precisely, and recorded with reliability (Baer et al.). Behaviors that are observable and measurable make this task possible. The *analysis* of the behavior requires the experimenter to confirm that the intervention was responsible for the change in behavior (i.e., a demonstration of control over the behavior; Baer et al.). Analysis is accomplished using data to monitor the progress of the behavior and to make decisions regarding the effectiveness of the intervention (Kearney, 2008).

Applied Behavior Analysis (ABA) is rooted in behavioral psychology theories and principles (as opposed to cognitive or developmental) to explain behavior (Alberto & Troutman, 2009). It proposes that behavior, both appropriate and inappropriate, is learned and influenced by the environment and subsequent consequences (Alberto & Troutman). For example, behavior that is followed by pleasant consequences is more likely to be repeated, whereas behavior followed by unpleasant consequences is less likely to be repeated.

Applied Behavior Analysis (ABA) is a process of applying these principles of behavior to change socially significant behavior to a socially significant degree (Alberto & Troutman, 2009). In other words, ABA focuses on behavior that is meaningful in the subject’s life and aims to change that behavior by more than a ‘statistical’ difference. A student may struggle with many academic issues in school, but an educational team may
elect to focus their efforts on the behaviors that will help that student function independently in the future. The team’s goal is to increase those skills to a level that can have a positive impact on the student’s life. It has also been argued that in order for research to meet the criteria of ABA, the chosen behavior must be socially important to those involved rather than because it is convenient to study (Baer et al., 1968).

*Antecedent and Consequential Considerations*

Because learning principles are the foundation of the behaviorist view of behavior, learning principles such as reinforcement, punishment, and antecedent strategies are involved when applying ABA (Alberto & Troutman, 2009). Certain considerations arise when working with these strategies.

*Antecedent Considerations*

Antecedents are events or conditions that happen prior to a behavior that may help identify why a behavior is occurring (Alberto & Troutman, 2009). Some antecedents are neutral (i.e., no affect on behavior) whereas others signal that a particular behavior is likely to be reinforced or punished (Kearney, 2008). For example, a student may scream out loud each time a teacher gives a command to the class or another student may fall asleep every day in second period after not sleeping the night before. Both students are being influenced by events that occur either directly before or more distantly to the targeted behavior. With the current impetus of Positive Behavior Supports, educators are encouraged to examine their learning environments in order to identify antecedents that promote safe, supportive conditions for all students (Alberto & Troutman).
Consequential Considerations

Events that follow a behavior (i.e., consequences) with regularity tend to affect how often that behavior occurs (Kearney, 2008). Examining consequences may help identify why a behavior is occurring. Many consequential considerations stem from operant conditioning learning principles such as reinforcement, punishment, and extinction (Alberto & Troutman, 2009). For example, one student may act out in class because she or he is reinforced by the rest of the class laughing, while another student acts out because they are reinforced by “getting a rise” out of the teacher. One student may cease to act out because of a phone call home (e.g., punishment), whereas another student may cease to act out if the teacher ignores the disruptive behavior (e.g., extinction). All instances illustrate the power of events that occur directly after a target behavior. These consequences may help explain (and be the reason) why the students are continuing their disruptive behavior.

Functional Assessment

As discussed above, educators are interested in understanding why certain behaviors occur and recognize that it is influenced by both what happens before and after the event. Behaviorists alike are concerned with which environmental factors increase, decrease, or maintain rates of behaviors (Skinner, 1953). A functional assessment is a method of gathering information in order to form a hypothesis on what variables may be influencing behavior (Alberto & Troutman, 2009). Functional assessments may be done by interviews with the student, teacher, and/or parents; checklists/rating scales completed by the student, teacher, and/or parents; or direct observation (Alberto & Troutman;
Kearney, 2008). Direct observations make use of the antecedent and consequential considerations previously mentioned by charting the events that occur directly before and after a target behavior in order to identify patterns.

Applied Behavior Analysis (ABA) has been recognized for introducing the idea that problem behavior may be functionally assessed in order to identify a social or communicative motive behind that behavior (Bambara, 2005). Functional assessments of behavior result in two motives (or functions): positive or negative reinforcement. Students exhibit behavior because they receive positive or negative reinforcement from the consequences of it. For example, some students may engage in behavior because it elicits attention from their teacher (i.e., something they are positively reinforced by). Other students may engage in behavior because it allows them to escape an aversive situation (i.e., negative reinforcement). Examining the function of behavior allows us to better understand why the behavior is occurring and thus, design the most appropriate way to intervene.

For that reason, researchers warn that interventions not solely focus on suppressing inappropriate behavior without attending to the function that behavior is serving. For instance, interventions that are matched to the function of a behavior are more likely to be effective compared to interventions that are only matched to the type of behavior (Watson & Steege, 2003). A particular intervention may stop a student from talking out during math class, but may not stop that student from replacing that behavior with throwing objects across the room (i.e., a different disruptive behavior that serves the
same function). Researchers recommend identifying the variables that control a behavior in order to design the most effective intervention (Watson & Steege).

History of ABA in Education

Researchers in the 1960s began to apply these principles of behavior to improve socially important behaviors. Efforts began to implement these strategies in the educational setting by adjusting the educational environment (Cooper et al., 2007). Early studies concluded that teacher attention successfully increased desirable classroom behavior (Hall, Lund, & Jackson, 1968), more effective methods of instruction and general classroom management strategies were designed (Keller, 1968; O’Leary & O’Leary, 1972), and the use of token economies to increase appropriate classroom behavior were reported (Kidder & Tague, 1965). Most studies focused their intervention to individual students rather than classwide applications.

Applied Behavior Analysis (ABA) in education has progressed to include more diverse populations and target behaviors. Sulzer-Azaroff and Gillat (1990) provided an analysis of 347 educational research articles published in the Journal of Applied Behavior Analysis between the years of 1968-1986. The analysis revealed that most of the students served were of elementary age (i.e., kindergarten through fourth or sixth grade), followed by special education students and preschool age students. Most target behaviors were related to classroom conduct (e.g., off-task, disruptive behavior, and aggression), then academic and language skills. Sulzer-Azaroff and Gillat noted that the focus on conduct fell during the second decade of research to be replaced by language and social skills
instruction demonstrating a new emphasis on skill development rather than on reduction of unwanted behavior.

Applied Behavior Analysis (ABA) can be credited for numerous effective teaching and behavior management interventions (Bambara, 2005) and is evident in classrooms today. Its focus on modifying behavior (through learning and reinforcement strategies) and analyzing data has become an integral part of education. ABA allows teachers to effectively modify students’ behavior and manage it positively. Many teachers today use reinforcement and punishment procedures, examine antecedent and consequential factors, and conduct functional assessments to manage behavior. ABA allows teachers to help their students master functional and academic skills in a systematic manner and gives teachers the tools to document and evaluate the progress (Alberto & Troutman, 2009). Behavior analysis strategies have produced results that are relevant to education and have encouraged a restructuring in educational environments (Sulzer-Azaroff & Gillat, 1990). Because of ABA, educators understand that environments can be manipulated to facilitate behavior change (Alberto & Troutman).

Positive Behavior Supports

Even more recently in education, tenets of ABA are recognized in Positive Behavior Supports (PBS). PBS emerged in the mid 1980s as an alternative to punishment/aversive behavioral strategies (Bambara, 2005). It works in conjunction with the request from The Individuals with Disabilities Education Act (IDEA, 2004) for more positive behavioral interventions by encouraging a more preventive and positive approach to problem behavior (Sugai & Horner, 2002a). PBS is defined as the broad
range of systematic and individualized interventions to achieve important social and learning outcomes while preventing problem behavior (Sugai et al., 2000). It aims to remediate deficits in the environment and an individual’s behavioral repertoire (E. G. Carr et al., 1999). Deficits in the environment may include lack of choice, inadequate teaching strategies, or poorly selected daily routines. Behavioral deficits may exist in communication, social skills, or self-monitoring (E. G. Carr et al.).

Positive Behavior Supports centers on behavioral science, practical interventions, social values, and a systems approach. Behavioral science refers to the idea that behavior is learned and impacted by its environment. PBS encourages individuals to evaluate the problem behavior (i.e., antecedents, consequences) in order to design practical interventions that teach more acceptable replacement behaviors. Interventions should be comprehensive, durable, and relevant in order to reflect social values. The systems approach encourages a team-based approach and refers to the idea that PBS can be implemented across the many systems that impact students (e.g., community, family, school, classrooms, non-classrooms; Sugai et al., 2000).

**Empirical Support**

The United States Department of Education, Office of Special Education Programs requested a review of the literature on PBS to obtain a general definition of the approach, an analysis of the database, an identification of gaps in knowledge, and suggestions for future use (E. G. Carr et al., 1999). Carr et al. responded with a meta-analysis on 109 published articles between the years of 1985-1996 which included 230 participants. The research reviewed found PBS being implemented by teachers and
parents; in schools or at home (one-third of the time); with elementary aged students with mental retardation exhibiting aggressive and/or self-injurious behavior. These outcome infer that PBS can be implemented by non-clinically trained individuals, in school settings, to improve serious behavior problems. It also shows the lack of research with adolescent populations, students without identified disabilities, and less severe behavior problems.

Overall, results indicated that (a) PBS is widely accepted by individuals with serious problem behavior, (b) assessment and the use of environmental interventions has become more popular, (c) PBS is effective in reducing the problem behavior of one half to two thirds of the cases reviewed (to an 80% criterion level), and (d) success rates double if a functional assessment of the behavior drives the design in intervention. It was also reported that PBS was equally effective whether individuals had multiple disabilities as opposed to a single disability, was more effective when the environment was modified and when the intervention was carried out by significant people in the individual’s life. Modest to substantial increases were reported in socially desirable behavior as well.

Another meta-analysis of the same database of information supported the results of E. G. Carr et al. (1999; Marquis et al., 2000). Marquis et al. concluded PBS effective across populations (i.e., age, gender, type of disability), problem behavior, environment, and type of intervention (i.e., stimulus vs. reinforcement based). They also concluded that implementing combinations of interventions (stimulus plus reinforcement based) were more effective than single interventions and that effectiveness was increased if assessment drove the choice in intervention.
Links to Applied Behavior Analysis

Positive Behavior Supports (PBS) highlights the use of data collection and analysis to inform decision making while focusing efforts on changing behavior to a socially significant level (Sugai et al., 2000). PBS procedures emphasize attention to and manipulation of antecedent events to reduce or prevent problem behavior and careful consideration of consequences to remove factors that maintain problem behavior (Sugai et al.). The core values of PBS—prevention of inappropriate and nonfunctional behaviors, application of research/evidence-based practice to build behavior repertoires, and creating contexts and processes that are person centered and supportive—were founded in ABA (Alberto & Troutman, 2009).

Continuum of Support

Positive Behavior Supports (PBS) is based on Walker et al.’s (1996) model and is intended to be maintained across three levels of support: (a) primary, (b) secondary, and (c) tertiary. Figure 1 illustrates a visual representation of the prevention model. Primary prevention is directed toward all students and focuses on preventing the development of problem behavior (Office of Special Education Programs [OSEP] National Technical Assistance Center on Positive Behavioral Interventions and Supports, 2009). It typically consists of rules, routines, and physical arrangements that are developed and taught by staff to prevent initial occurrences of problem behavior (OSEP). It is anticipated to be a sufficient level of support for approximately 80% of the school population.

Students identified as not responding to the primary intervention (approximately 15%) are then involved in more intense support (secondary level) in efforts to avoid more
serious behavior problems. Secondary prevention is designed to provide more targeted interventions to students at risk for developing more serious problem behavior (OSEP, 2009). This level of intervention is typically delivered to small groups of students or through simple individualized intervention plans (OSEP).

Students requiring even more intervention are involved in tertiary level interventions. Tertiary prevention is designed to focus on the needs of students who exhibit patterns of problem behavior (approximately 5% of school population; OSEP, 2009). The goal of this level of support is to decrease already existing problem behavior.
and increase alternative behaviors/skills (OSEP). It typically includes a functional behavioral assessment to determine the function of the problem behavior and to guide creation of behavioral interventions and behavior plans. This level focuses on the individual needs of the student and the unique characteristics associated with them in order to create an individualized behavior support plan (OSEP).

Overview of Antecedent Strategies

Literature exists to support the practice of arranging antecedent conditions to help develop desirable behaviors, to diminish problem behaviors, and to design successful environments (Cooper et al., 2007). Table 1 outlines some of the more recently investigated antecedent interventions. Antecedent strategies attempt to modify behavior by altering the availability of reinforcement (e.g., delivering or withholding reinforcement for a specific stimulus) or by modifying the antecedent event itself (e.g., precorrection; Cooper et al., 2007). Teachers commonly utilize antecedent strategies such as assigning seats, establishing procedures, and enforcing rules to create safe classrooms; changing seating arrangements to prevent misbehavior; and creating activities to avoid misbehavior during transition times.

Overview of Consequential Strategies

A table illustrating both punishment and reinforcement (i.e., consequential) behavioral strategies typically implemented in the educational setting is provided in Appendix A.
**Punishment Procedures**

Punishment, by definition, results in a decrease in behavior following the presentation of an aversive event or removal of a favorable event (Kazdin, 1977, 2001).

Table 1

**Antecedent Interventions**

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Study</th>
<th>Subjects</th>
<th>Target Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice of instructional procedures</td>
<td>Daly, Garbacz, Olson, Persampieri, &amp; Ni (2006)</td>
<td>Two middle school students with Behavioral Disorders</td>
<td>Reading fluency</td>
</tr>
<tr>
<td>Choice opportunities to change clothes</td>
<td>Carlson, Luiselli, Slyman, &amp; Markowski (2008)</td>
<td>Two children on the PDD Spectrum</td>
<td>Public disrobing and urinary incontinence</td>
</tr>
<tr>
<td>Intensive early, positive, proactive interventions</td>
<td>Cartledge, Singh, &amp; Gibson (2008)</td>
<td>---</td>
<td>General academic and behavior problems</td>
</tr>
<tr>
<td>Public posting of rules &amp; teacher movement</td>
<td>Musser, Bray, Kehle, &amp; Jenson (2001)</td>
<td>Three elementary students with social and emotional disorders</td>
<td>Compliance</td>
</tr>
<tr>
<td>Corrective academic prompts</td>
<td>Ebanks &amp; Fisher (2003)</td>
<td>19-year-old man with mental retardation and PDD</td>
<td>Destructive behavior (e.g., SIB, aggression toward others, property destruction)</td>
</tr>
<tr>
<td>Flooding</td>
<td>Saigh, Yule, &amp; Inamdar (1996)</td>
<td>---</td>
<td>PTSD in school aged children</td>
</tr>
<tr>
<td>Noncontingent reinforcement</td>
<td>Kahng, Iwata, Thompson, &amp; Hanley (2000)</td>
<td>Three adults with developmental disabilities</td>
<td>SIB and aggression</td>
</tr>
<tr>
<td>Prompts</td>
<td>Manley, Collins, Stenhoff, &amp; Kleinert (2008)</td>
<td>Three elementary students with cognitive disabilities</td>
<td>Telephone skills</td>
</tr>
</tbody>
</table>
Overcorrection and response cost strategies are considered punishment because they aim to decrease disruptive behavior by removing positive incentives. Although punitive practices have become common in our schools (Cartledge et al., 2008), evidence suggests that punitive procedures may result in negative side effects (i.e., increases in problem behavior) and are not likely to develop positive behaviors in individuals (Kazdin, 1977) or create safe and orderly schools (Horner, Sugai, Todd, & Lewis-Palmer, 2005).

Researchers argue that punishment procedures are best utilized as support to a reinforcement procedure and should be reserved for situations that involve decreasing a physically dangerous behavior, when Differential Reinforcement procedures cannot be successfully implemented, or to temporarily suppress a problem behavior while reinforcing a positive one (Kazdin, 2001). Research warns that there is no evidence to strongly support punishment procedures over other techniques (Kazdin) and those undesirable behaviors can be eliminated by reinforcement rather than punishment (Donnellan & LaVigna, 1986). Some researchers argue that punishment should never be used (Donnellan & LaVigna, 1990).

Limitations

Punishment procedures have been criticized for being limited in their effectiveness and ability to promote meaningful outcomes (Bambara, 2005). They are considered reactive, consequence-based, and short-term focused (Bambara). Although punishment procedures reduce inappropriate behavior, the effects are short-lived (Bambara), have been associated with negative side effects (i.e., increases in problem
behavior), and are not likely to develop positive behaviors in individuals (Kazdin, 1977) or create safe and orderly schools (Horner et al., 2005).

Furthermore, others argue that punishment procedures do not address the needs of today’s schools (Skiba, 2002) and may create a false sense of security (Sugai & Horner, 2002b). A fear exists that if school is aversive, more problem behavior may surface (Horner et al., 2005). Many have argued for a more positive proactive approach in dealing with problem behaviors in schools (Sugai & Horner). Examples are found at the Center for Study and Prevention of School Violence, the Center for Positive Behavioral Interventions and Supports, and the Office of Special Education Programs.

Reinforcement Procedures

Whereas punishment procedures aim to decrease behavior, positive reinforcement systems aim to increase behavior. Positive reinforcement has been proven to eliminate a wide range of problem behaviors and has been used to develop and increase socially appropriate behaviors (Kazdin, 2001). Kazdin argued that positive reinforcement “represents the core feature of behavior change programs in applied settings” (p. 199). Although Differential Reinforcement of Other Behavior (DRO) contains both a reinforcement and extinction component, it is considered a positive reinforcement procedure because the goal of the intervention is to increase appropriate behavior through reinforcing it.

Types of Differential Reinforcement

Differential reinforcement (DR) grew out of the theory that we expect to see more of a behavior that is reinforced (positively or negatively) rather than behavior that is not.
It is founded on the principle of reinforcing desirable behavior and ignoring undesirable behavior. Thus, DR is composed of both reinforcement and extinction components.

*Differential reinforcement of alternative behavior (DRA).* DRA procedures reinforce alternative behaviors to the undesirable target behavior (Kazdin, 2001) while extinguishing the undesirable ones (Miltenberger, 2001). In theory, by increasing occurrences of the alternative behavior, occurrences of the target behavior decrease. A more “appropriate” alternative behavior is reinforced while the “inappropriate” behavior is ignored. This technique is most effective when the chosen alternative behavior serves the same function as the behavior being replaced (Alberto & Troutman, 2009). For example, DRA may be implemented with a student who frequently disrupts/interrupts the teacher while she is lecturing the class in order to gain attention from her. The teacher will choose an alternative behavior (e.g., on-task) to reinforce and will ignore or interrupt the disruptive behavior. Each time the target behavior occurs, the behavior is either ignored or interrupted, and the student is redirected to the alternative behavior, and thus, reinforced for it. In this case, each time the student is on-task they will be rewarded with positive teacher attention (e.g., labeled praise) and disruptive behavior will be ignored.

Miltenberger (2001) stated that the following three concepts will help determine if DRA is the appropriate intervention to implement: (a) the goal of the intervention should be to increase a predetermined positive behavior, (b) the alternative behavior must already be in the student’s repertoire, and (c) a powerful reinforcer must be available. Additionally, seven steps have been identified in enhancing the success of a DRA intervention: (a) define the desirable behavior/the behavior that will be reinforced, (b)
define the undesirable behavior, (c) identify the reinforcer, (d) reinforce the desirable behavior immediately and consistently, (e) eliminate reinforcement for the undesirable behavior, (f) use intermittent reinforcement to maintain the target behavior, and (g) program for generalization (Miltenberger).

*Differential reinforcement of incompatible behavior (DRI).* DRI is a variation of DRA in which the alternative behavior to be reinforced is physically incompatible with the target behavior (Kazdin, 2001; Miltenberger, 2001). The incompatible behavior is typically the direct opposite of the target behavior, which by definition means they cannot occur at the same time (Kazdin; Miltenberger). With DRI, reinforcing the incompatible behavior decreases the occurrence of the target behavior. For example, a student may be reinforced for folding their hands in their lap rather than hitting another student. Repp and Deitz (1979) suggested collecting baseline data on the disruptive behavior as well as the incompatible behavior. This gives the researcher a realistic estimate of the occurrence of the behaviors in order to create goals for the intervention. As with DRA, choosing an incompatible behavior that already exists in the student’s repertoire enhances the intervention’s success.

*Differential reinforcement of low rates of behavior (DRL).* DRL interventions reinforce an individual for decreasing the occurrence of a target behavior. DRL is similar to a changing criterion design (Alberto & Troutman, 2009) in that reinforcement is delivered if the target behavior decreases to a criterion level (Miltenberger, 2001). The DRL procedure is best utilized when a low rate of the target behavior can be tolerated or the target behavior is maladaptive only because of its high rate of occurrence.
Criterion levels may be modified to achieve the appropriate level of the behavior (Kazdin, 2001). DRL interventions are advantageous because they can reduce behavior while still reinforcing and do not require drastic changes in behavior to be deemed successful (Alberto & Troutman). For example, a DRL may be implemented to encourage a dominating student to allow other students to participate more in class. If baseline data indicate the student is contributing to the class discussion 20 times during a class period, the teacher can set the initial goal at 15 times per class. In this situation, the student is reinforced for contributing to the class discussion only 15 times as opposed to 20 times. The teacher continues to set the goals successively lower until an appropriate level of contributions is reached.

Miltenberger (2001) compared the two ways DRL can be implemented. Full-session DRL requires the individual to achieve a specific criterion for an entire session (e.g., math class, lunch time, morning session) in order to be reinforced. For example, a student may be required to raise their hand only 3 times during math class.

Spaced-responding DRL requires an individual to allow a specific amount of time (i.e., “space”) between responses. In this case, the individual is required to wait 10 minutes before raising his or her hand again.

Differential reinforcement of other behavior (DRO). DRO reinforces the individual contingent on the absence or omission of the target behavior (Reynolds, 1961). DRO involves both reinforcing the omission of the target behavior and an extinction component (i.e., withholding reinforcement for occurrences of the target behavior; Mazaleski, Iwata, Vollmer, Zarcone, & Smith, 1993). Although the name of the
procedure hints that other behaviors are reinforced, it is actually the absence of the target behavior that results in reinforcement. In other words, the individual is reinforced only when she or he does not exhibit the target behavior (e.g., an individual is reinforced for not talking about during a specific interval of time). Some refer to this form of DR as the differential reinforcement of zero rates of behavior or omission of behavior (Alberto & Troutman, 2009). Theoretically, the problem behavior decreases if periods of time without the behavior increase (Miltenberger, 2001). Miltenberger argued that DRO interventions are most effective when the reinforcer for the target behavior can be identified and eliminated, when the interval of reinforcement is based on the baseline rate of behavior, and when the reinforcement is delivered for the absence of behavior for the entire interval of time (whole interval DRO).

DRO has been found useful for reducing high rates of disruptive behavior (Kazdin, 2001). Four steps have been identified by Miltenberger (2001) in order to implement DRO with success: (a) identify the reinforcer for the problem behavior, (b) identify the reinforcer to use as part of DRO, (c) choose the initial DRO interval of time, and (d) eliminate the reinforcer for the problem behavior and deliver the reinforcer for the absence of the problem behavior.

To implement this technique, data are collected on the occurrence of the target behavior and the average duration between behaviors is determined (i.e., the inter-response interval). This information drives the reinforcement schedule that is delivered if the target behavior does not occur. For example, if data gathered during baseline suggest that the target behavior occurs once every minute, reinforcement will be
delivered at the end of a minute if the target behavior has not been exhibited. Once behavioral control has been established, the interval required for delivery of the reinforcement is progressively increased (de Zubicaray & Clair, 1998).

Overview of DRO

DRO became popular in the late 1960s and 1970s because of its simplicity, focus on reinforcement (rather than punishment), and ethical foundation. At this point in time in behavioral research, most attention was focused on punishment procedures because they were proven effective (Repp, Deitz, & Deitz, 1976). Although effective, typical punishment procedures (e.g., overcorrection, time out, aversive stimulation) were beginning to lose popularity due to the severity of their consequences (Repp et al.) and education’s movement toward more preventive models. Researchers argued that punishment may not be ethical and overcorrection may not be realistic (Poling & Ryan, 1982). Furthermore, the use of punishment alone fails to actively teach a positive alternative behavior. Procedures such as satiation, extinction, and differential reinforcement began to emerge.

Theoretically, the problem behavior decreases if periods of time without the behavior increase (Miltenberger, 2001). Miltenberger argued that DRO interventions are most effective when the reinforcer for the target behavior can be identified and eliminated, when the interval of reinforcement is based on the baseline rate of behavior, and when the reinforcement is delivered for the absence of behavior for the entire interval of time (whole interval DRO). Other researchers highlighted the effectiveness of utilizing
strong, relevant reinforcers to strengthen the omission of the target behavior (Mazaleski et al., 1993).

DRO has proven both effective and ineffective with different populations and behaviors (see Homer & Peterson, 1980; Poling & Ryan, 1982), and the DRO literature encompasses a variety of broad categories. The following review focuses on research conducted with preschool and elementary aged children that illustrates (a) research supporting DRO as an effective intervention, (b) research concluding DRO as an ineffective intervention, (c) factors to consider when assessing the efficacy of DRO, and (d) recommendations for future research.

Research Supporting DRO

Since the 1970s, very little research has examined DRO as a single component intervention (i.e., as opposed to pairing it with other interventions as packages). Unfortunately, these studies offer a more direct indication of the power of DRO in changing behavior as no other intervention components can be deemed responsible.

One such study compared DRO and noncontingent reinforcement as control procedures in increasing a preschool child’s compliance (Goetz, Holmberg, & LeBlanc, 1975). In the noncontingent condition, the subject received sporadic verbal praise for compliance consisting of unrelated pleasantries (e.g., “you’re wearing a pretty dress today”). The noncontingent condition also involved sporadic occasions of the teacher walking within three feet of the subject for non-compliance. During the DRO condition, if the subject did not comply within 30 seconds, the teacher moved within three feet of her and delivered a verbal comment. This condition was characterized by more
occurrences of teacher presence than the noncontingent condition. Verbal comments for noncompliance during this condition were unrelated to reinforcement (e.g., “I don’t blame you for not picking up the toys”).

Results indicated that the DRO procedure decreased compliance faster and in fewer sessions than did noncontingent procedures. DRO produced a more immediate and dramatic decrease in the target behavior. These results need to be considered with caution because the spirit of DRO was not evident in this research. DRO is meant to reinforce other behavior. If the target behavior in this study was compliance, the DRO condition should have reinforced behavior other than compliance which was not evident. The three conditions were also weakly defined. It was unclear how teacher presence and verbal comments were used throughout all three conditions.

Repp et al. (1976) implemented DRO on a one-to-one basis (i.e., individually) with three mentally retarded children. Subjects’ target behaviors varied among hair twirling, hand biting, and thumb sucking and reinforcement varied between teacher praise and candy. Inappropriate responding was reduced substantially and was attributed to the DRO schedule as demonstrated by the multiple baseline and ABAB designs chosen.

In another study examining DRO as a single-component intervention, Daddario, Anhalt, and Barton (2007) examined the effectiveness of DRO in reducing the disruptive behavior of typically developing preschool children during circle time. Disruptive behavior was defined as touching a peer or teacher with hands, fingers, arms, or feet. It included poking, tackling, hitting, and kicking the teacher or peers. The study incorporated an event recording system to monitor the disruptive behavior across the
baseline and treatment conditions. During baseline, the classroom teacher tallied instances of disruptive behavior on paper, while during treatment the teacher set aside an M&M candy after observing each instance of disruptive behavior (in order to calculate a total at the end of the activity). During intervention, the teacher set a kitchen timer to the appropriate interval time (as determined by taking the average number of occurrences as derived from the baseline data) and conducted circle time. When the timer beeped, the teacher delivered M&Ms (and labeled praise) to randomly selected students for not emitting disruptive behavior, and reset the timer. If disruptive behavior occurred at any point during the interval, the teacher ignored the disruptive behavior, set aside an M&M, and reset the timer.

Although results indicated success in reducing disruptive behavior, several limitations were noted. The chosen AB design (preferred among teachers) does not illustrate experimental control of the intervention and therefore does not allow solid conclusions regarding the strength of the intervention in changing behavior. Variability in the data during treatment was observed and may be attributed to the low treatment integrity exhibited by the classroom teacher. Low treatment integrity scores also indicated that the teacher may not have followed the DRO procedure as designed and therefore, the intervention may not be solely responsible for the behavior change observed. Suggestions for future research include increasing experimental control by implementing DRO using a different design (e.g., reversal or multiple baseline), increasing treatment integrity by providing performance feedback in addition to
checklists, and implementing the intervention with different populations and for different target behaviors.

*Implemented as a Multi-Component Intervention*

Researchers have argued that DRO can be made more effective by combining it with other procedures (Homer & Peterson, 1980). It is commonly implemented as a multi-component intervention with a token economy. This seems a natural combination due to the fact that by definition, DRO must reinforce behavior, and token economies have been proven effective across individuals, settings, and behaviors (see Kazdin, 1977). Repp et al. (1976) reported success in reducing levels of inappropriate responding in a classroom using a DRO schedule of reinforcement paired with a token economy. Three students with high levels of “talk outs” were targeted in this experiment. Each student earned tokens if no talk outs were observed. Tokens could then be exchanged for games, activities, and refreshments.

DRO was successfully paired with a token economy with three individuals with moderate mental retardation in a large self-contained classroom (Poling, Miller, Nelson, & Ryan, 1978). Target behaviors included leaving assigned seat, hitting others, throwing objects, putting hands near mouth, and drooling. An ABA single subject research design was used. During baseline, teachers recorded the occurrence of the target behaviors using frequency counts and no instructions were given to the participants. Tokens were distributed for the non-occurrence of the target behavior and could be exchanged for food, social interaction, or access to toys. Verbal praise was paired with the delivery of each reinforcer.
Results indicated a dramatic decrease in “talking out” during intervention and an increase during the reversal phase. Therefore, the reduction in responding can be attributed to the DRO intervention. The authors reviewed several advantages of using DRO and reported that it represents a potentially valuable tool for behavior management in the classroom. Despite these positive findings, some limitations were observed. Although interval times were increased with improved behavior, determination of the initial interval time was not addressed. This violates the spirit of DRO in that baseline behaviors directly establish the reinforcement schedule (Repp, Deitz, & Speir, 1974).

In a related study, five multi-handicapped children were successful in reducing their disruptive behavior using a DRO with token economy intervention (Didden, de Moor, & Bruyns, 1997). Each participant chose their own reinforcement which was categorized as a back-up reinforcer (i.e., a token was exchanged for a reinforcer). Students received a token during the intervention phase for refraining from their target behavior for the entire interval of time. Prior to each session, the teacher explained the rules of the token system and which behaviors were appropriate or not appropriate. Each participant chose one reinforcer to be earned during each treatment session.

Results indicated that the DRO-tokens intervention, using reliably assessed back-up reinforcers, was effective in reducing the children’s disruptive behaviors to near zero levels in the classroom. The importance of appropriate back-up reinforcers was highlighted in this study. Although positive outcomes were reported, several questions arose. The authors collected data using a partial interval system. Frequency counts give a more accurate account of behavior (Alberto & Troutman, 2009). Also, the participants
were limited in earning only three tokens per session. This aspect of the study limited the subjects’ ability to earn more tokens which could then be exchanged for more back-up reinforcers. Typically token economies include a wide range of back-up reinforcers to provide incentive to earn higher amounts of tokens in exchange for more highly valued items or events (Kazdin, 1977).

Three students with autism were taught to self-manage a DRO program in order to reduce their disruptive behavior in a related research study conducted by Newman, Tuntigian, Ryan, and Reinecke (1997). Target behaviors were out of seat and nail flicking. Students were given praise and a token at the end of each interval for which they did not engage in the target behavior (i.e., whole interval DRO); tokens were then traded in for favored consumables or break time. The timer was reset for each observation of a target behavior and a brief explanation given (e.g., “you’re out of your seat, I have to reset the timer”). Although the study focused more on the effects of the self-managing component, DRO as part of a multi-component intervention proved to reduce disruptive behaviors across all conditions and to maintain these results at follow-up.

In another study investigating the effectiveness of DRO as a multi-component intervention, Deitz, Repp, and Deitz (1976) conducted two experiments examining the effectiveness of DRO in reducing disruptive behavior of mentally retarded students in the classroom. The teacher was the primary data collector and data collection procedures in both experiments were identical. The first student, a 13-year-old male, was able to earn one point for each interval of time in which he did not emit a disruptive response. His responding decreased immediately and remained low throughout treatment conditions.
The second experiment compared two sets of students who excessively talked out during class. The students were awarded a half minute of free time at the end of the class period for each two-minute interval during which they did not emit talking out behavior. Results for the second experiment indicated reduced levels of talking out. The teacher indicated that the intervention was easy to implement and was satisfied with the results of both experiments.

The results of this study illustrated the importance of modifying the DRO protocol in order to meet the needs of the teacher. A partial interval recording system is typically not a preferred method of data collection in this situation due to its tendency of overestimating occurrences of behavior. Even so, it was used to make the intervention more feasible for the teacher and the effectiveness of the intervention remained clear. This simple modification offers support in direct contradiction to Vollmer, Iwata, Zarcone, Smith, and Mazaleski’s (1993) claim that DRO is a “cumbersome [intervention] to administer over long periods of time . . . because it requires continuous monitoring of a client’s behavior” (p. 10).

DRO has also proven effective when implemented with interventions other than token economies. DRO was implemented in three separate experiments (i.e., DRO alone, DRO with reprimands, and DRO plus extinction) in a study by Rolider and Van Houten (1984) designed to decrease inappropriate behavior. Each experiment had one typically developing participant ranging in age from 4.5 to 12. The first experiment, a reversal research design, concluded that DRO alone was ineffective in reducing the frequency of abusive behavior a 4.5 year-old girl directed at her baby sister. The addition of
reprimands to the DRO protocol resulted in declines in abusive behavior. The second experiment found DRO plus a reprimand procedure to reduce more bedtime tantrums than a DRO plus extinction contingency. The third experiment, a multiple baseline research design, implemented DRO using a momentary time sampling approach. The subject’s parents looked to see if the target behavior was occurring every 3 minutes. This experiment concluded that DRO plus reprimands reduced thumb-sucking more effectively than DRO alone.

Overall results concluded that DRO plus reprimands produced greater reductions in target behaviors compared to DRO alone and DRO plus extinction. Nonetheless, a type of DRO was proven effective at decreasing critical target behaviors such as physical aggression toward a baby and refusing to sleep in one’s own bed.

Results of this study should be interpreted with caution for several reasons. For example, Experiment 1 (i.e., abusive behaviors toward a baby) implemented DRO using a random interval time and did not conduct a reinforcement assessment. Thus, one cannot be certain that more robust results may have occurred during the DRO alone condition or that the participant was motivated to change her behavior. The results of Experiment 3 (i.e., thumb-sucking) reported that DRO plus reprimand produced a greater decrease in the target behavior, but the DRO alone condition also produced promising results. In this experiment, DRO was also implemented using a momentary time sampling procedure. Research has proven that whole interval DRO is more effective in initially reducing target behaviors while momentary DRO is more effective at maintaining those results (Barton, Brulle, & Repp, 1986; Repp, Barton, & Brulle, 1983).
It is also important to consider that all three experiments reported more success in the DRO plus reprimands conditions, but the confounding variable of the reprimands was not controlled for. In other words, the reprimands looked different across the three experiments. In the first experiment, the mother held her daughter by the shoulders while making eye contact, told her to stop the behavior, and gave her a reason why. In the second experiment, the parents were instructed to approach the child and say, “look at me.” Then they pointed their index finger at the child’s chest, told her to stop crying/asking to get out of bed, told her she was bothering them and to go back to bed. In the third experiment, the mother was told to approach her son and say, “look at me,” point her finger and say, “stop sucking your thumb and take your finger out of your mouth.” Therefore, it is difficult to tease out whether it was the DRO intervention or the manner in which the reprimands were given that impacted behavior.

Shabani, Wilder, and Flood (2001) implemented an intervention package including discrimination, DRO, and self-monitoring with a 12-year-old boy with autism, attention deficit/hyperactivity disorder, mild mental retardation, and seizure disorder. Prior to the start of the study, a reinforcement assessment and functional assessment were conducted. The reinforcement assessment established powerful reinforcers while the functional assessment (of rocking) concluded that the subject engaged in body rocking regardless of what was happening in the immediate environment.

Results revealed the treatment package successful in decreasing the stereotypic behavior of rocking (i.e., the dependent variable) in the classroom environment. Although positive results were obtained, causal conclusions cannot be made between DRO and the
reduction of inappropriate behavior due to the multi-component nature of the intervention. Whereas one can conclude that the package as a whole reduces disruptive behavior, one cannot readily conclude that DRO alone decreases behavior. Another flaw in the study stemmed from the trend of the baseline data. In their graph, Shabani et al. (2001) reported a decreasing trend of rocking behavior during baseline. Assumptions cannot be made that it was the intervention on its own that was responsible for the decrease in behavior because the behavior was headed in that direction before intervention (Cooper et al., 2007).

Concerns About Existing Methodology

Studies comparing DRO to response cost (RC) have yielded results in favor of the latter. Sindelar, Honsaker, and Jenkins (1982) conducted two case studies in which DRO and RC were implemented to increase the attending behavior of two distractible students. In the first study, a 7-year-old girl with a learning disability was given half-page segments of stories to read while in her resource room. During the DRO condition, she was able to earn tokens for each half page read without looking away. During the RC condition, she was given 12 tokens to begin and lost one token for the first look away on any half-page segment. Although the DRO component produced significant decreases in off-task behavior (i.e., looking away), the data remained variable. The RC component’s results were more stable and produced better results.

In the second study, a 10-year-old girl diagnosed as behaviorally disordered was involved in an ABC design (i.e., baseline, DRO, RC) during her math lesson in the regular classroom. During the DRO condition, the student was able to earn tokens on a
variable interval schedule for not looking away. During the RC condition, the student was given 15 tokens to begin and lost one token each time she looked away during the session. Tokens were counted on an abacus for both conditions. Data collected during the DRO condition indicated that the student’s look aways were not decreasing and subsequently might be increasing due to her looking at the abacus.

Although the authors concluded the RC condition as successful in reducing look aways to “well below baseline levels” (p. 9), visual inspection of the graph indicates a large portion of overlapping data between conditions. The results of this study may be interpreted as somewhat limited because the authors did not use baseline data to determine interval times in the DRO condition. Additionally, the DRO and RC conditions were not similar regarding the location of the abacus. Data revealed an increase in look aways during the DRO condition due to the student paying more attention to the abacus than the reading passage. Therefore, the researchers placed the abacus on the floor for the RC condition. This introduced a confounding variable not addressed in the study. Despite the limitations associated with these studies, they can lead one to conclude that DRO was effective in reducing the first student’s off-task behavior and might have proven successful in experiment two had the abacus been relocated earlier.

Conducting a similar comparison, Conyers et al. (2004) used an alternating treatment design to evaluate the effectiveness of DRO and response cost in a preschool classroom setting. Response cost and DRO sessions occurred on alternating days and disruptive behavior was observed at the classwide level. Students were instructed on how
the procedure would be implemented prior to beginning the intervention phase. Students earned tokens which could be exchanged for candy at the end of the 15 minute session.

As expected, Conyers et al. (2004) found that the response cost component resulted in more drastic and lasting reductions of disruptive behavior compared to the DRO component. Most behavioral literature supports the conclusion that reductive consequences (i.e., response cost) reduces behavior more immediately than positive reinforcement (i.e., DRO). However, the DRO component was an effective approach to decreasing disruptive behavior as well as reinforcing appropriate behavior.

Results of this study should be considered with caution for several reasons. The interval times were selected randomly by the research assistants. This practice goes against the “spirit” of DRO which is meant for reinforcement to be delivered according to the rate of disruptive behavior (Miltenberger, 2001). Furthermore, feedback was given to the students during the response cost condition and not during the DRO condition. This introduces a confounding variable. One cannot conclude that it was the response cost intervention itself (i.e., loss of tokens) resulting in decreased behavior or the combination of the intervention and corrective feedback. Although the results reported more drastic reductions in behavior during the response cost condition, the DRO condition also produced promising results (i.e., a 37% change in disruptive behavior from baseline to the DRO treatment phase).

Summary

Debate continues on the effectiveness of DRO throughout the literature. Early studies concluded pessimistic results (e.g., Corte, Wolf, & Locke, 1971; Foxx & Azrin,
1973) while others reported the reverse (e.g., Poling et al., 1978; Repp et al., 1976). Both the Sindelar et al. (1982) and Conyers et al. (2004) studies concluded RC as the more robust intervention over DRO. This led researchers to examine the DRO protocol in hopes of identifying specific explanations for the evident discrepancy. Several factors have been identified as critical to the protocol yet frequently overlooked in the literature. Some argue that when these issues are included in a DRO intervention, stronger outcomes may result (Repp, Felce, & Barton, 1991).

Factors to Consider When Assessing the Efficacy of DRO

Reinforcement

One theory responding to the question of why DRO may be ineffective revolves around the idea of reinforcement. In general, researchers agree that reinforcer potency is vital to successfully changing behavior (Mason, McGee, Farmer-Dougan, & Risley, 1989). Specific to DRO, researchers have identified three components in addressing the reinforcement component of the intervention: (a) identify the function of the target behavior, (b) magnitude of the chosen reinforcement, and (c) variation (Mazaleski et al., 1993). Mazaleski et al. suggested identifying the function of the target behavior in order for it to be extinguished during the extinction component of DRO. One must identify the function in order to stop reinforcing it. The reinforcement chosen (as part of the reinforcement component of DRO) must be powerful enough to compete with the already successful reinforcing inappropriate behavior and should be varied in order to avoid satiation (Poling & Ryan, 1982; Repp et al., 1983; Repp et al., 1991).
The importance of back up reinforcers has been proposed as a possibility of DRO being deemed ineffective (Didden et al., 1997). Their study ensured that a reinforcer assessment was conducted prior to implementation in order to guarantee that the reinforcers were in fact reinforcing to the students. Prior to the study, each student was individually assessed for reinforcement choices which resulted in five different rankings of preferred reinforcement that would be used as back-up reinforcers. Students earned tokens for refraining from target behaviors and exchanged them at the end of each session. Results showed that the DRO intervention using tokens was “highly effective in decreasing disruptive behavior of all participants” (p. 65).

This action highlights a common theme in applied behavior analysis in that behavior will not change if the reinforcement is not robust enough. The Didden et al. (1997) study emphasized that DRO may result in more powerful outcomes if the reinforcement aspect is addressed. Studies that maintain DRO as an ineffective intervention need to assure that the reinforcement offered for the change in behavior was truly reinforcing rather than the researcher’s poor choice of reinforcement (Repp et al., 1983; Vollmer et al., 1993).

**Momentary DRO vs. Whole DRO**

In the research literature, DRO has been implemented in two ways: whole interval DRO (wDRO) and momentary DRO (mDRO). The wDRO procedure has been defined as administering the reinforcer if behavior other than the target behavior is exhibited for the entire interval. In contrast, the mDRO procedure involves administering the reinforcer if the other behavior is exhibited at the time of a prompt (e.g., the sound of a timer; Repp et
Researchers have speculated that the two types of DRO schedules may be associated with the effectiveness of the intervention (Repp et al., 1991).

Repp et al. (1983) conducted two studies to compare the effectiveness of wDRO and mDRO. Results indicated that wDRO decreased inappropriate behavior more than mDRO. In other words, wDRO was more effective at initially reducing the behavior, but mDRO may be effective at maintaining the level of behavior. In fact, Barton et al. (1986) argued that mDRO can be effective in maintaining a behavior change once a wDRO schedule has initially reduced it. Researchers have noted the difficulty of implementing wDRO accurately due to the constant attention required by the observer. As a result, mDRO was valued due to the reduced amount of time teachers/intervention agents invest (Barton et al.).

Success was reported by Conyers, Miltenberger, Romaniuk, Kopp, and Himle (2003) in implementing DRO at the classwide level with 22 preschoolers exhibiting disruptive behaviors. The researchers used a reversal design with multiple treatment conditions to measure the effectiveness of momentary DRO (mDRO) and whole DRO (wDRO) with various reinforcers. Treatment conditions were complex, but involved subjects receiving stars and praise for not exhibiting disruptive behavior. Students that had earned a pre-determined amount of stars during the session were permitted to exchange them for a small toy or surprise edible reinforcer.

The authors reported a decrease in the frequency of disruptive behaviors exhibited by students. Also, the number of children who engaged in disruptive behavior decreased after the DRO intervention was implemented. Conyers et al. (2003) reported that wDRO
with edible reinforcement produced the greatest decreases in disruptive behavior. The authors also reported that mDRO produced modest results when paired with tangible reinforcers.

Limitations of this study include randomly assigned interval times and a lack of consistency of variables. DRO is meant to utilize inter-response times determined by the rate of behavior and duration of time between behaviors during baseline (Repp et al., 1974). If this formula is not followed, one cannot conclude that the intervention was effective or not. Also, this study examined the use of edible reinforcement with wDRO and tangible reinforcement with both wDRO and mDRO, but did not address edible reinforcement with mDRO.

**Intervals of Time**

The length of the interval in a DRO intervention is the amount of time between reinforcement (i.e., interresponse time) and has been calculated in different ways. Researchers have suggested using learning performance in determining interresponse time (IRT; Cooper et al., 2007). Intervals calculated using learner performance, rather than arbitrarily chosen intervals, will be more sensitive to change in the participant’s behavior resulting in a more accurate reinforcement schedule. This also enables a practitioner to lengthen the interval as the behavior improves (Cooper et al.).

Some researchers have argued that the interval should represent the mean interval between responses during baseline. For example, Repp et al. (1974) proposed a formula and set the initial interval to the inverse of the mean response rate during baseline. In
other words, if students are disruptive 5 times (on average during baseline) during a 15-minute class period, the initial interval would be set at 3 minutes.

Another option in determining the IRT is averaging the previous sessions. Vollmer et al. (1993) found DRO to be highly effective in reducing self-injury when averaging the previous three or five (respectively) sessions for their participants. Though the Vollmer et al. study compared DRO to non-contingent reinforcement, the importance of determining the IRT according to past research was evident. Lindberg, Iwata, Kahng, and DeLeon (1999) averaged the previous three sessions to obtain IRT for their participants. The initial IRT was determined by averaging the baseline data and changed daily as a result of the previous day’s performance. Although past research was not directly cited in this study’s method section, the authors computed IRT according to recommendations of past research (e.g., Repp et al., 1974).

Researchers have hypothesized that selecting inappropriate interval times can produce ineffective results (Repp et al., 1976; Repp et al., 1991). More specifically, intervals that were too large tended to be less successful influencing the target behavior. Repp et al. (1991) hypothesized that interval size may impact the effectiveness of DRO and therefore, conducted two experiments testing the idea. One classroom determined their reinforcement schedule by calculating the mean of the last 3 days of baseline data. The other classroom doubled the mean. Both schedules were successful in reducing inappropriate behavior, but the shorter interval (using the mean) was “considerably more effective” (p. 421). The second experiment tested the same hypothesis but across students in different settings. Again results showed the shorter time interval (using the mean) to be
more effective. Both experiments highlight the conclusion that interval size is relevant in ensuring the effectiveness of DRO and that intervals should be calculated in relation to the baseline data.

In other words, for DRO to be most effective, the other behavior needs to be reinforced before the target behavior occurs. Studies that have not determined interval size in this manner may have mistakenly concluded DRO as ineffective (e.g., Conyers et al., 2004; Sindelar et al., 1982). Accurate collection of data during baseline facilitates this piece of the intervention.

*Interval Reset vs. Continuous Interval*

Although not typically included in discussions debating the effectiveness of DRO, the subject of whether the timer continues to its selected endpoint or resets to zero when a target behavior occurs still remains an important decision in implementing DRO. Two options are available when the target behavior is emitted during a DRO intervention: The timer can be reset immediately (i.e., interval reset), or the interval can end as scheduled (wait until the timer beeps to reset; continuous interval). Researchers have reported success with both the continuous interval (Lindberg et al., 1999; Repp et al., 1983) and interval reset (Daddario et al., 2007; Newman et al., 1997; Repp et al., 1976; Vollmer et al., 1993) options.

Interval reset has been the customary option for DRO (Repp et al., 1983). In this option, the timer is immediately reset upon occurrence of the target behavior. This tends to promote an easy connection for children to see between emitting the target behavior and not receiving reinforcement. For example, students realize that when they emit the
target disruptive behavior, the timer is reset and no reinforcement is given. In other cases, it allows the interventionist the opportunity to correct inappropriate behavior and point out that reinforcement will not be delivered until the inappropriate behavior is not emitted. For example, the interventionist may say “you will not earn a sticker until you have kept your hands to yourself.” Regrettably, this may lead to accidental reinforcement of unwanted behavior (i.e., the student may be reinforced by the teacher providing corrective feedback or by the attention given in resetting the timer; Foxx, 1982).

The continuous interval option (i.e., waiting for the timer to exhaust before resetting) allows the interventionist to ignore the target behavior (i.e., extinction) without accidentally reinforcing it. This option provides more consistent interval lengths for interventionists to work with, but may not be as effective with some populations of students. For example, by allowing the timer to continue until exhaustion (rather than resetting it immediately) a child’s ability to connect the reinforcement to their behavior may be limited.

Summary of Existing Research and Methodology

DRO has been proven effective in decreasing talk outs (Repp et al., 1976); leaving assigned seats, hitting others, and throwing objects (Poling et al., 1978); aggressive behavior (Rolider & Van Houten, 1984); and stereotypic behavior (Shabani et al., 2001), among others. It has mainly been implemented with students with disabilities (e.g., autism, Newman et al., 1997; mental retardation, Deitz et al., 1976; Poling et al., 1978; and multiple handicaps, Didden et al., 1997). It has been successfully implemented on its own (Daddario et al., 2007); paired with token economies (Didden et al., 1997;
Poling et al., 1978; Repp et al., 1976); and self-monitoring techniques (Shabani et al.,
2001). DRO has been successfully implemented using single subject research designs
(i.e., reversal, multiple baseline; Didden et al., 1997; Rolider & Van Houten, 1984).

Researchers have concluded that factors such as reinforcement, momentary versus
whole implementation, and interval size can impact the overall effectiveness of DRO.
Some agree that the reinforcement component of DRO should be powerful enough to
change the disruptive behavior (Poling & Ryan, 1982; Repp et al., 1983; Repp et al.,
1991). Whole interval DRO has been proven more effective in initially reducing
disruptive behavior while momentary DRO maintains that level of behavior (Barton et
al., 1986). Finally, interval size should be related to the rate of behavior during baseline
conditions (Repp et al., 1991).

Questions remain in the literature regarding the effectiveness of DRO
implemented with young children, typically developing populations, and when
reinforcement, whole vs. momentary interval, and interval times are targeted, controlled,
and systematically explored.

Future Directions

Results regarding the effectiveness of DRO have been inconsistent as a result of
irregularity in implementation. Some researchers overlook the critical issues of
reinforcement (Conyers et al., 2003; Mazaleski et al., 1993), momentary versus whole
interval implementation (Repp et al., 1983), and calculating interval times; and therefore,
may mistakenly conclude DRO as unsuccessful. It is for this reason that studies need to
address these issues with integrity in order to make valid conclusions about the intervention.

It may also be beneficial in the DRO literature for the intervention to be expanded to include younger students and students without disabilities. Measures of social validity, treatment acceptability, and treatment integrity can be included to examine the value, ease, and opinions of the intervention agents and the degree with which the intervention was implemented as designed. Of particular interest is the effect of a pre-treatment acceptability (and intervention effectiveness) discussion on a teacher’s treatment acceptability of an intervention (Elliott, 1988). Also critical to future research are the inclusion of treatment integrity measures (Gresham, Gansle, & Noell, 1993) and the impact of performance feedback on treatment integrity (Daddario et al., 2007). It may also be valuable for DRO to be investigated as a single component intervention (as opposed to being paired with a token economy; Shabani et al., 2001).

The current study addresses the issues of reinforcement, momentary versus whole DRO, and determining interval length to disqualify these possible confounding variables. It aims to extend the literature on the effectiveness of DRO as an isolated intervention rather than as part of a multi-component intervention (e.g., DRO with a token economy). It also aims to extend the work of Conyers et al. (2003) and Daddario et al. (2007) in applying DRO at the classwide level with young children. In doing so, this study examines the effectiveness of implementing whole interval DRO (wDRO) during treatment and momentary interval DRO (mDRO) during maintenance, at the classwide level to reduce talking out in typically developing children using a multiple baseline
design across settings. Measures of social validity and treatment integrity are also utilized.

**Research Questions**

1. Will DRO with a preference assessment component and implemented utilizing the average IRT from the previous five days of data decrease talking out to near zero levels?
2. Will whole DRO decrease levels of talking out and momentary DRO maintain those levels during the maintenance phase?
3. Does performance feedback coupled with a review of student outcome data maintain acceptable treatment integrity levels?
4. What is the acceptability level of a class wide application of DRO for decreasing talking out behavior in a kindergarten setting?
5. Is there a difference in the perceived intensity of disruptive behavior after implementation of class-wide DRO?
CHAPTER II
METHODOLOGY

Participants and Setting

The participants in this study were typically developing kindergarten children that emitted typical forms of disruptive behavior (e.g., talking out, poking, pushing) rather than more severe behavior (e.g., self injurious behavior). This study was conducted in an urban, tuition-free charter school in northeastern Ohio. The school housed grades kindergarten through eighth grade and had a total student population of approximately 300 students. The school served children from low and middle socioeconomic status. The majority of the students received free or reduced lunch. The participating classroom had 24 typically developed kindergarten students. Typical classroom activities included teacher led instruction, floor time, individual work, gym, lunch, and recess. This was the classroom teacher’s first year of teaching.

The participants were gathered by contacting the school’s psychologist. The school psychologist put the researcher in contact with the principal, who contacted White Hat Management, to receive permission for the study. A brief summary of the proposal was provided to the principal along with other supporting documentation (e.g., a similar study run by the researcher with preschool children, the researcher’s curriculum vita). After initial permission was granted, the classroom teacher whose class most closely met the requirements described above was contacted by the principal and asked to participate.
Subsequently, the principal contacted the researcher and classroom teacher to confirm participation and provide introductions.

The teacher and each student’s parent or guardian was asked to sign consent forms (for both the target class and social comparison class). Participation was voluntary and could have been terminated at any time. Each student was asked to assent to participation in the intervention as well. Copies of each of these forms can be found in the appendices (Appendix B). Consent forms were sent home with each of the students. Parents that expressed questions or concerns were contacted by the researcher (over the phone) to provide clarifications. The principal was also given copies of each of the parent letters (e.g., consent form and ECBI explanation).

Procedure

Preparation of the Classroom

Before onset of the study, the researcher met with the teacher to arrange the classroom in a way that facilitated video recording, to discuss the target behaviors, and to choose the settings. The video camera was set up in a manner to allow complete viewing of the chosen setting. It was also set up in advance to allow time for the students to become familiar with it and desensitized to it. Masking tape was used to designate where the video was set up each day to allow for consistency. Target behaviors (i.e., the dependent variable) were chosen by the teacher based on social significance. In other words, the teacher chose the behaviors that were most disruptive to the educational setting. The target behavior was operationally defined and the settings were chosen in consultation with the researcher and teacher. The teacher identified floor time and math
as the two most challenging times of day to control the target behavior. Floor time was chosen as the first setting to receive the intervention because it was the most problematic for her. Behavior rating scales (SESBI-R and ECBI) were distributed to the teacher and parents. The researcher met with the teacher of the social comparison classroom to discuss her class’s role in the project. Settings were chosen to closely match the target classroom, placement of and practice with the video camera were addressed, and consent forms were collected.

Reinforcement Pre-Assessment

Reinforcer potency is vital to successfully changing behavior (Mason et al., 1989). Therefore, this study applied a reinforcer assessment package which required the teacher to determine a pool of classroom items that were reinforcing to the entire class and then offer the choice of a reinforcer before each session during the study. Before the study began, the researcher presented the class with 10 different items that had been identified as reinforcers. Reinforcement options were limited to quickly consumable items such as edibles (e.g., Goldfish® crackers, pretzels, M&Ms®, Skittles®, Smarties®, Fruit Loops®, or Cheerios®); tangibles such as stickers or stamps; or social rewards such as high fives or “class snaps.” As the researcher presented each item, she asked the class to raise their hands if they “would like to earn [that item] for good behavior in school.” The researcher tallied how many of the students raised their hands for each item. Items that were agreed upon by at least 80% of the class were selected as the pool of reinforcers for the study. This was a slight variation of the Mason et al. model which implemented the technique individually to three children with autism.
Data Collection

Data collection began at the onset of each activity and continued for 20 minutes of the activity. Data collectors started their stop watches when the teacher said, “Let’s begin ____” and stopped their watches when the teacher said, “____ is finished.” Tallies (frequency count) of talking out were written along with the length of time (in minutes and seconds) of each activity in order to calculate the rate of behavior. This was ultimately reported as response per minute as a unit of analysis. Appendix C contains the standard data collection sheet. Data collection continued in the same manner during all phases of the study.

The target behavior (i.e., talking out) was defined as any student talking or making sounds to the teacher or another student without raising his or her hand and without being called on by the teacher. Talking out also included “off topic” conversations. The talking out needed to be at a voice level loud enough for the observer to hear (whispering and mumbling did not count). Talking out examples included repeating what another student said, grunting, making an “ooh” sound, and self-talk louder than conversation level during seat work. Non examples included when students responded to a rhetorical question posed by the teacher (e.g., “Should I set my timer?”), whispering or low level mumbling, quiet self-talk during individual seat work, when the teacher allowed the students to answer as a group/without raising their hand (even though it fits the target behavior definition). Each occurrence (rather than the amount of what was said) was tallied (event recording). For example, one student making an “ooh” noise
counted as one occurrence, whereas one student who answered a question without raising his or her hand also counted as an occurrence.

The number of talk outs emitted by all students during an activity was divided by the total duration of the activity in order to provide a measure of the rate of the target behavior (responses per minute). Interresponse time (IRT) was calculated by taking the inverse of that rate. The primary data collector’s data were used to calculate the IRT. The rate was calculated for each day of baseline, then averaged, and inversed in order to determine the initial interval (IRT) for the first week of intervention. The researcher was responsible for calculating the IRT and reporting it to the teacher. The teacher’s kitchen timer was set according to the inverse of the mean rate of responding (IRT) during previous sessions. This rate was rounded (when necessary) to the nearest half-minute to increase ease of implementation. Rounded rates were always in favor of the students (i.e., shorter intervals for increased opportunity for reinforcement). Initial intervention phase IRT was determined by baseline data while subsequent IRT was calculated using the average IRT across the previous 5 days of intervention. IRT was calculated separately for each setting.

Data were collected through a video camera and posted to a private website blog. This allowed the data collectors to view the video at their convenience and multiple times (if necessary). The researcher was responsible for turning on (and off) the video recorder and for collecting the video and posting it to the blog. The teacher signaled the beginning and end of each activity (and thus the beginning and end of data collection) by introducing/ending the activity (e.g., “Let’s begin floor time;” “Floor time is finished”).
Pre-baseline footage was used to test the appropriate placement of the camera, picture clarity, audio level, and student desensitization.

Two experienced graduate students assisted with data collection (one primary researcher and one as reliability check). A meeting was held between the researcher and graduate students to discuss the target behavior definition and method of data collection before the study was initiated. Pre-baseline video provided training footage and both students were trained to at least 80% inter-rater agreement. During training, the data collectors viewed a 15-minute video (through the website blog) and, together, tallied the dependent variable according to the behavior definition outlined. Agreement was calculated, but was not used to determine reliability. Afterwards, the data collectors independently viewed a second 15-minute video (through the website blog) and tallied the dependent variable. Agreement was greater than 80%, thus, a third video was not used.

During the study, one graduate student served as the primary data collector and the second observer collected data on randomly assigned days (directed by the primary researcher) in order to co-code at least 30% of the baseline and treatment data (approximately every third day). Each graduate student viewed the video independent of the other. Interobserver agreement (IOA) was calculated by dividing the smaller number of responses by the larger number and multiplying by 100 (Kazdin, 1982). If interobserver agreement dropped below 80% at any time during the study, the two graduate students were in contact with the researcher to review behavioral definitions in order to avoid observer drift and other compromises to the reliability of the data.
Single Subject Research Design

A multiple baseline across settings design (i.e., floor time and math) was implemented in order to analyze the effects of the intervention in two settings. The multiple baseline design allowed the researcher to simultaneously analyze more than one setting (Alberto & Troutman, 2009). First, in a multiple baseline across settings, baseline data were collected in each setting. Second, the intervention was implemented in the first setting when baseline data indicated stability across all settings (Kazdin, 1982), while data continued to be collected in the remaining setting. Third, intervention was introduced for the second setting when the first setting showed a positive trend or criterion had been established. The intervention remained in effect for the first setting and data continued to be collected for the remaining setting. This sequence was continued until the intervention had been introduced to both settings.

Experimental control is illustrated at each point the intervention is introduced across settings. Each setting should change when the IV is introduced. Unique to the multiple baseline design, variables should have a discrete relationship (i.e., IV should only affect one variable at a time). Therefore, a functional relationship only exists if changes are apparent in the specific setting being manipulated by the IV. Settings that have not yet received intervention should remain at baseline levels (i.e., changes in the remaining settings should not be observed; Kazdin, 1982).

The current study’s design improved on the AB design implemented in the Daddario et al. (2007) study by increasing the value of experimental control and external validity (generalizability). The study began by collecting baseline data in both of the two
settings (i.e., floor time and math). The whole interval DRO intervention was introduced during floor time when baseline data indicate stability across both settings. Baseline data continued to be collected on the remaining setting (i.e., math). The intervention was introduced during math when the data during floor time showed a positive trend (3 data points demonstrating decreases in talking out).

Internal validity was controlled for by utilizing two graduate assistants as data collectors and retraining them on the behavioral definition when IOA fell below 80%. Plans were established to retrain the teacher if treatment integrity levels fell below 80%. Fortunately, this did not occur.

Social Validity/Social Comparison

Social comparison involves comparing the target individual to “normal” functioning peers (Kazdin, 2001), by asking if the intervention brought the target behavior to a level equal to that of peers who are functioning appropriately (Kazdin, 1982). In order to determine if the intervention brought the talking out to a level equal to that with peers, data were collected one day per week (during baseline, intervention, and maintenance) in the other kindergarten classroom. The other kindergarten class was videotaped in two settings (floor time and math) and talking out was tallied according to the same procedure described above.

Baseline

The teacher was instructed to signal the beginning (and ending) of the activity and conduct the session as usual. Baseline data were collected (in both of the chosen settings) for at least 5 days and until data stability was achieved. Baseline data continued to be
collected during math (i.e., the second setting) while the intervention was implemented
during floor time. The number of talk outs emitted by all students during an activity was
divided by the total duration of the activity in order to provide a measure of the rate of the
target behavior. The rate was calculated for each day of baseline and then averaged in
order to set the initial interval for the first week of intervention. The kitchen timer was set
according to the inverse of the mean rate of responding during baseline.

Pre-Intervention

Research suggests providing teachers with information about the intervention
(regarding its effectiveness; Von Brock & Elliott, 1987) and clearly specifying the
intervention procedures and guidelines (so it is implemented as originally designed; Lane,
Bocian, MacMillan, & Gresham, 2004) before implementation. Therefore, the researcher
and teacher met to discuss and role play the DRO protocol (Appendix D). Topics
addressed during the meeting included discussing the proven effectiveness of the DRO
intervention in the research literature and the importance of treatment integrity. At the
same time, training was provided on delivering labeled praise.

Research also states that performance feedback coupled with social support and
evidence of student outcomes increases levels of treatment integrity (Witt, Noell,
LaFleur, & Mortenson, 1997). For that reason, the researcher modeled the whole interval
DRO intervention and provided performance feedback on the teacher’s simulated
implementation. The treatment integrity checklist, performance feedback, and student
outcome data aspect were also discussed. Specifically, the researcher informed the
teacher that every few days during the study, the researcher would watch the video to
ensure adherence to the DRO protocol. Performance feedback was provided the following day and the students’ progress was shared via an Excel graph. The intervention was not implemented until the teacher’s simulated performance yielded treatment integrity scores above 90%. The initial social validity measures, the BIRS (Von Brock & Elliot, 1987) and GAS (Sladeczek, Elliott, Kratochwill, Robertson-Mjaanes, & Stoiber, 2001), were completed at this time. After examining the baseline data, the teacher and researcher established the final goal of the intervention, as measured by the GAS, at 2 talk outs per minute. SESBI-R and ECBI behavior rating scales were collected as well.

*Intervention/Whole DRO*

*Setting one.* The teacher asked the class to “pick one” of the two reinforcers offered. Once a reinforcer was agreed upon by the students, the teacher signaled the beginning of the activity, started the timer, and began the lesson. When the timer beeped, the teacher delivered reinforcement and labeled praise (e.g., “Good job keeping your hands to yourself”) to the entire class if all students did not talk out for the entire interval (whole interval DRO; group contingency) and reset the timer. Both the chosen reinforcer and the labeled praise functioned as reinforcement. If talking out occurred during the interval, the teacher ignored the inappropriate behavior and reset the timer (interval reset application). No reinforcement was distributed.

On the seventh day of intervention, the DRO protocol was modified to increase teacher feasibility by allowing her to post icons (i.e., tokens) rather than delivering reinforcement at the end of each interval (i.e., the wDRO with icons phase). Each icon represented one earned reinforcement, with students receiving their reinforcement at the
Conclusion of the lesson. This allowed the teacher’s lesson to continue with fewer
distractions. After five more days of intervention, the reinforcement system was modified
in order for each student to choose his or her own reinforcement rather than the entire
class agreeing on one (i.e., the wDRO with icons plus individual reinforcement phase).
The teacher continued to give the class a choice of two reinforcements before beginning
the lesson, but the researcher wrote down which student chose which reinforcement. The
intervention phase in setting one continued for approximately five weeks.

Setting two. Intervention in the second setting began utilizing the wDRO with
icons (i.e., tokens) and individual reinforcement modification from the first setting. The
teacher offered the class a choice of reinforcement; however, the researcher recorded
each student’s preference. Then, the teacher signaled the beginning of the activity, started
the timer, and began the lesson. If the entire class did not talk out for the entire interval,
she posted an icon and gave labeled praise. If talking out occurred during the interval, the
teacher ignored the inappropriate behavior, reset the timer, and did not distribute
reinforcement. The intervention phase in setting two continued for one week.

Pre-Maintenance Meeting

Similar to the procedure followed in the pre-intervention meeting with the
teacher, the researcher met with the teacher before implementing the maintenance phase
(momentary DRO) to discuss and role play the DRO protocol (Appendix E). The
researcher modeled the momentary interval DRO intervention and provided performance
feedback on the teacher’s simulated implementation. The treatment integrity checklist,
performance feedback, and student outcome data components were also discussed as
described above. The intervention was not implemented until the teacher’s simulated performance yielded treatment integrity scores above 90%. The GAS (i.e., social validity measure; Sladeczek et al., 2001) was completed at this time.

**Maintenance/Momentary DRO/Reversal**

*Setting one.* For the initial interval, the teacher was asked to set the kitchen timer to the same interval as the last day of intervention. For subsequent intervals, she used the inverse of the mean rate of the previous five days. When the timer beeped, the teacher posted an icon on the board and delivered labeled praise to the entire class if all students did not talk out at the sound of the beep (i.e., momentary interval DRO using group contingency) and reset the timer. In other words, all talking out during the interval was ignored unless it occurred at the sound of the beep and thus resulted in no reinforcement. The maintenance phase continued for approximately one week resulting in a total of six weeks of intervention implementation.

*Setting two.* In an effort to further enhance feasibility, a reversal phase (i.e., return to baseline) was implemented in the second setting rather than the momentary application of DRO. Therefore, the intervention was withdrawn and the teacher returned to implementing her math lessons as usual.

**Post-Maintenance Meeting**

At the conclusion of the study, a final meeting was held between the researcher and teacher to discuss and evaluate the effectiveness of both the whole interval and momentary interval DRO intervention, complete the final social validity measures (i.e., GAS & BIRS), and distribute/collect the remaining SESBI-R and ECBI rating scales.
Measures

Social Validity/Treatment Acceptability

Treatment acceptability has been defined by Kazdin (1981) as “judgments by laypersons, clients, and others of whether treatment procedures are appropriate, fair, and reasonable for the problem or client” (p. 493). Researchers have developed models that hypothesize a fundamental link between treatment acceptability, treatment use, treatment integrity, and treatment effectiveness (Witt & Elliott, 1985). In other words, interventions that are not “acceptable” by those implementing them will be less likely to be implemented as designed (with integrity) and, consequently, will be less effective. Research also recommends assessing the teacher’s perceived acceptability and effectiveness of an intervention before beginning implementation (Von Brock & Elliot, 1987).

Teacher acceptability of procedures and results (i.e., the intervention) was measured using the Behavioral Intervention Rating Scale (BIRS; Von Brock & Elliot, 1987; Appendix F) at the beginning and the end of the study. The BIRS consists of 24 self-report items rated on a 6-point Likert scale (1 = strongly disagree; 6 = strongly agree). Items examine the acceptability of the intervention (e.g., “Most teachers would find this intervention suitable for the problem behavior described,” “The intervention is reasonable for the behavior problem described”), the effectiveness of the intervention (e.g., “The intervention should produce enough improvement in the child’s behavior so the behavior is no longer a problem in the classroom”), and the intervention’s time to
effect (e.g., “The intervention would quickly improve the child’s behavior”). Higher mean level scores indicate higher acceptability.

**Treatment Integrity**

Treatment integrity “refers to the extent to which an intervention is implemented as originally designed” (Lane et al., 2004, p. 37). When interventions are not implemented accurately the internal and external validity can be jeopardized (Gresham et al., 1993). In other words, one cannot conclude that the intervention was responsible for the change in behavior and opportunities for further replication may be difficult (Lane et al., 2004). Researchers agree that because the outcome of school-based interventions is so critical (improved academic performance), measures need to be taken to insure that interventions are being implemented accurately (Lane et al.). Per procedures outlined by Telzrow and Beebe (2002), this study incorporated the following elements to enhance treatment integrity: (a) choose an intervention that is based on research, (b) create and utilize a treatment integrity checklist, (c) provide the interventionist with training and feedback, and (d) conduct frequent integrity checks.

A treatment integrity checklist (Appendixes D and E) was used to provide an index of the teacher’s adherence to the DRO protocol. The checklist recorded the occurrence and nonoccurrence of each intervention component and a percentage was calculated (Gresham et al., 1993). The teacher completed the checklist after each activity in which DRO was implemented as a self-monitoring measure. Additionally, the researcher viewed the video tape, evaluated the teacher’s performance every fourth day of intervention, provided performance feedback (i.e., in person and informally after
school), and shared current student outcome data in order to increase adherence to the protocol. It was established that the teacher would be retrained by the researcher if treatment integrity data revealed performance less than 80% adherence to the protocol (omitting two components), but this did not occur.

Social Validity of Goals

Wolf (1978) defined social validity through three levels: social significance of the goals, social acceptability of the procedures used, and social importance of the results. This study utilized the Goal Attainment Scaling (GAS; Kiresuk & Sherman, 1968) to assess the social validity of the treatment goals. The objective of the GAS is to operationally define the target behavior(s) and final goal of the intervention and then rate the behavior according to a point scale. The final goal is operationalized via a range of positive to negative outcomes (from +2 to -2) that will provide the criterion for successful completion of the intervention.

The GAS outcome behaviors are to be seen as “perceived best guesses as to what behaviors can be expected” (R. A. Carr, 1979, p. 89). Specifying the level of outcome can be reported as frequency, percent, or intensity (Kiresuk, Smith, & Cardillo, 1994). Because the GAS has been implemented in many fields such as education, rehabilitation, medicine, and social work (Kiresuk et al.) the criterion for allocating a behavior outcome to a specific attainment level (e.g., +2, +1, 0, -1, -2) has been calculated in various ways. In the current study, criterion levels were calculated in consultation with the researcher and teacher and reflected expected, yet realistic changes in behavior. For example, the 0 attainment level was anchored at baseline with the best possible outcome anchored to +2
and the worst possible outcome at -2 (Sladeczek et al., 2001). For this study, a +2 represented a 75% positive change in behavior whereas a -2 represented a 75% negative change in behavior. A +1 and -1 reflected a 50% change in behavior. In other words, if baseline data indicated 20 talk outs, +2 reflected a 75% decrease in talk outs (i.e., 5 talk outs) and a -2 reflected a 75% increase in talk outs (i.e., 35). The researcher rated the class behavior every 2 weeks during the study and reported that progress to the teacher (Appendix G).

*Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R)*

The SESBI-R is a behavior rating scale that measures the severity and extent to which conduct problems in school settings are considered problems. The measure is composed of 38 items. Teachers report a frequency score (1 = *Never*, 7 = *Always*) and report whether they perceive the behavior to be a problem. The SESBI-R yields two scores: an Intensity score (how frequently the behavior occurs) and a Problem score (if the behavior is considered a problem). Eyberg and Pincus (1999) have suggested T-scores above 60 as cutoffs to estimate the severity of child behavior problems at school. The SESBI-R scales have shown internal consistency coefficients (Cronbach’s alpha) of 0.98 (Intensity Scale) and 0.93 (Problem Scale; Querido & Eyberg, 2003). The teacher was asked to complete a SESBI-R for each student before and after the intervention was implemented to provide pre/post measures of problem behavior.

*Eyberg Child Behavior Inventory (ECBI)*

Similar to the SESBI-R, the ECBI is a behavior rating scale that measures the severity and extent to which conduct problems in the *home* are considered problems. The
measure is composed of 36 items. Parents rate the frequency of disruptive behavior on a scale (1 = *Never*, 7 = *Always*) and whether they perceive the behavior to be a problem. The ECBI scales have shown internal consistency coefficients (Cronbach’s alpha) ≥ 0.97 (Intensity Scale) and ≥ 0.95 (Problem Scale; Eyberg & Pincus, 1999). Eyberg and Pincus have suggested T-scores above 132 on the Intensity Scale and 15 on the Problem Scale as cutoffs to estimate the severity of child behavior problems at home. This measure was also completed before and after the study to provide pre/post measures of problem behavior.

**Data Analysis**

Data were reported as a rate of behavior and presented in line graphs. Results were evaluated through visual analysis (the traditional approach of data interpretation in single subject research; Parsonson & Baer, 1978). Visual analysis has been the preferred method of data analysis in applied settings for several reasons (Tawney & Gast, 1984). It reveals strong and socially significant results which are what teachers are interested in (Alberto & Troutman, 2009). It has also been argued as a way of providing strong variables and interventions. Weak variables (i.e., ones too insignificant to be identified through visual analysis) will typically stay in the laboratory (Parsonson & Baer) rather than being successful in applied settings. Therefore, interventions that show success through visual analysis are powerful (Baer, 1977). It allows a researcher (or teacher) to evaluate individuals or small groups on an on-going basis and make decisions about the intervention throughout the study (Tawney & Gast). This encourages more individualized instruction and implementation of interventions (Tawney & Gast).
Specifically, changes in mean, immediacy of change, trend levels, percent nonoverlapping data, variability, and number of data points were examined. A change in mean referred to a change in the average instances of a behavior between phases (Kazdin, 1982). The data from each phase of the study (i.e., baseline, intervention, maintenance) were averaged (i.e., a mean was calculated) in order to compare the changes. Immediacy of change was determined by evaluating the “jump in behavior” immediately after the intervention had been introduced. Specifically, the last datum point during baseline was compared to the first datum point during intervention, and the last datum point during intervention was compared to the first datum point during maintenance. Changes in trend referred to the change in the slope of the data (Kazdin). Slopes may be labeled as ascending, descending, or no slope. Trends during baseline are expected to move in a counter-therapeutic direction, whereas trends during intervention should lean toward a therapeutic direction. In particular, the slopes during each of the phases were calculated using Microsoft Excel, labeled, and evaluated according to the direction they were headed.

Scruggs, Mastropieri, and Casto (1987) argued that the proportion of overlapping data (between treatment and baseline conditions) can be an important evaluative criterion in determining the effectiveness of an intervention. Researchers hope to find a large amount of data points, and therefore a large percentage of data points, that do not overlap (i.e., percent nonoverlapping data; Tawney & Gast, 1984). In other words, the amount of behavior emitted during baseline should not be the same as the amount of behavior emitted during treatment. Percent of nonoverlapping data (PND) is calculated by
indicating “the number of treatment data points that exceeds the highest baseline data point in the expected direction and dividing by the total number of data points in the treatment phases” (Scruggs et al., 1987, p. 27) and multiplying by 100 (Tawney & Gast). Table 2 outlines the framework for determining effectiveness as described by Scruggs and Mastropieri (1998).

Data variability was determined by examining the fluctuations in data. Variability was calculated for each phase by dividing the mean in half and adding and subtracting the quotient to the mean to establish the upper and lower boundaries. Per the recommendations of Alberto and Troutman (2009), points that fell outside the boundaries rendered the data variable. A minimum of three data points per condition is recommended in order to accurately determine the level of stability and trend of the data.

Table 2

_PND Intervention Effectiveness Ratings_

<table>
<thead>
<tr>
<th>PND score</th>
<th>Intervention considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above 90%</td>
<td>Very effective</td>
</tr>
<tr>
<td>70-90%</td>
<td>Effective</td>
</tr>
<tr>
<td>50-70%</td>
<td>Questionable</td>
</tr>
<tr>
<td>Below 50%</td>
<td>Ineffective</td>
</tr>
</tbody>
</table>

(Tawney & Gast, 1984). Therefore, the number of data points in each phase were counted in order to adhere to this standard.
Descriptive statistics were reported as associated with the SESBI-R and ECBI. A dependent $t$-test for paired samples statistical analysis was used because the sample was dependent (i.e., the same sample was tested twice: pre-test, post-test; Fraenkel & Wallen, 2006). A 2-tailed was performed because the research question stated that a difference (not in a specific direction) was the desired outcome (Fraenkel & Wallen).
CHAPTER III

RESULTS

The purpose of this study was to investigate the effectiveness of whole interval DRO (during intervention) and momentary interval DRO (during maintenance) on the frequency of talk outs in the classroom. Chapter 3 provides results that answer the following research questions:

1. Will DRO with a preference assessment component and implemented utilizing the average IRT from the previous five days of data decrease talking out to near zero levels?

2. Will whole DRO decrease levels of talking out and momentary DRO maintain those levels during the maintenance phase?

3. Does performance feedback coupled with a review of student outcome data maintain acceptable treatment integrity levels?

4. What is the acceptability level of a class wide application of DRO for decreasing talking out behavior in a kindergarten setting?

5. Is there a difference in the perceived intensity of disruptive behavior after implementation of class-wide DRO?

Visual Analysis

Following the traditional approach of data interpretation in single subject research (Parsonson & Baer, 1978), the intervention’s impact on the rate of disruptive behavior
(i.e., talking out) was evaluated through visual analysis to answer the first 2 research questions. Specifically, changes in mean, immediacy of change, trend levels, percent nonoverlapping data, variability, and number of data points were examined.

**Setting One**

As depicted in Figure 2, DRO did not reduce talking out behavior to near zero levels during floor time. Seven days of baseline data collected during floor time (i.e., the first setting to receive the intervention) resulted in an ascending trend and a mean rate of 4.3 talk-outs per minute (range = 1.6-7.5). These data illustrate the likelihood that without intervention, the talk-outs would remain at this level or increase. Twenty days of intervention data collection resulted in a slightly descending trend with a mean response rate of 3.2 talk-outs per minute (range = 1.8-5.2). A slight immediacy of change was observed after the intervention was implemented in the first setting. One hundred percent of the data between baseline and treatment overlapped. According to criteria outlined by Scruggs and Mastropieri (1998; see Table 2), these data represent an ineffective intervention. Per the recommendations of Alberto and Troutman (2009), variability was calculated for each phase by dividing the mean in half and adding and subtracting the quotient to the mean to establish the upper and lower boundaries. Subsequently, points (across all phases) fell outside the variability boundaries which render all data variable. Finally, seven days of maintenance data were collected resulting in an ascending trend and a mean response rate of 2.2 talk-outs per minute (range = 0.6-2.7).
Figure 2. Rate of talking out across baseline, treatment, and maintenance phases for both settings.
Setting Two

Similarly, Figure 2 illustrates that DRO did not reduce talking out to near zero levels in the math setting (i.e., setting two). Twenty three days of baseline data collected during math time resulted in a descending trend and a mean response rate of 3.9 talk-outs per minute (range = 1.9-10.6). A visual inspection of five days of intervention data reveals a descending trend with a mean response rate of 2.2 talk-outs per minute (range = 1.5-3.5). Sixty percent of the data between baseline and treatment overlapped. Data in both the baseline and treatment phases were deemed variable, whereas data in the reversal phase was not. However, upon visual inspection, intervention data appear less variable than baseline data. Seven days of reversal data coincided with the intervention results (i.e., descending trend, mean response rate of 2.2 talk-outs per minute; range = 1.4-2.6).

Across Settings

Similarities were noted when examining the data across both settings. The target behavior, talking out, was observed to be increasing at an alarming rate during the first six days of data collection (i.e., baseline phase for the first setting) in both settings with average responding rates as high as 6.1 responses per minute during math and 4.3 during floor. Outliers like this were not observed again during the study. Most importantly, the intervention effectively reduced the talking out in both settings with average responding rates during floor reduced from 4.3 rpm to 3.2 rpm, and 3.9 rpm to 2.2 rpm during math. Data were variable throughout baseline and treatment phases across both settings, with the intervention resulting in a descending trend in both floor and math. Another similarity
was noted in the results of the intervention. The implementation of the maintenance phase in the first setting produced the same changes in behavior ($M = 2.2$) as noted during the intervention phase in the second setting ($M = 2.2$).

Differences in the data across settings were noted in trend levels during baseline and immediacy of change when the intervention was implemented. Baseline behavior during floor time resulted in an ascending trend, while baseline behavior during math represented a descending trend. Also, immediacy of change was observed when the intervention was first implemented during floor. Although a small immediacy of change was detected during math, it was in an anti-therapeutic direction.

**Treatment Integrity**

In order to address research question 3, treatment integrity was assessed through checklists and video review. The teacher was instructed to complete a treatment integrity checklist (Appendices D and E) after each setting during which the intervention was implemented. Ultimately, teacher checklists were completed for 24 out of 32 sessions, which accounted for 75% of the treatment and maintenance sessions. It was noted by the researcher that 7 of the checklists (the entirety of the maintenance phase checklists) were completed during the post-study meeting. The teacher self-reported 100% adherence to the intervention protocol. Additionally, the researcher completed the same treatment integrity checklist after observing the teacher implementing the intervention through video on the blog. The researcher’s data accounted for approximately 25% of the treatment sessions during floor (the first setting), 29% in the maintenance phase during floor, and 20% of the treatment sessions during math (the second setting). The researcher
also reported a 100% adherence to the intervention protocol. Due to the high rates of treatment fidelity, there was no occasion to implement the performance feedback component after the initial meeting.

Social Validity Measures

Acceptability

Research question 4 was answered by assessing the teacher’s acceptability of the procedures and results of the intervention, before and after the study, as measured using the Behavioral Intervention Rating Scale (BIRS; Von Brock & Elliot, 1987; Appendix F). Factor analysis of the BIRS generates three factors: Acceptability, Effectiveness, and Time to Effect (Elliott & Von Brock Treuting, 1991). Overall treatment acceptability ratings obtained by the teacher prior to and at the end of the intervention were in the below acceptable range ($M = 2.7$). The mean pre-intervention BIRS rating was 3.9 and the mean post-intervention rating was 1.4. Six was the highest possible mean rating that could be obtained via the BIRS. Pre-intervention scores indicated a slightly higher than neutral agreement with the treatment components, whereas the post-intervention scores indicated a general strong disagreement with the treatment components. Overall Effectiveness ratings were in the below acceptable range ($M = 2.1$) and overall Time to Effect ratings were also in the below acceptable range ($M = 2.3$). The teacher’s rating of the Effectiveness and Time to Effect decreased from pre-intervention ($M = 3.1, 3.5$ respectively) to post-intervention ($M = 1.1, 1$ respectively).
Goals

Social validity of goals was measured using the Goal Attainment Scale (GAS) during each treatment integrity feedback meeting which resulted in 3 data points. Directly upon implementation of the intervention, a GAS score of +1 was obtained and maintained throughout the study in both settings. This score indicates that DRO produced a 50% change in talking out wherein a 75% change was the study’s ultimate goal.

Comparison Data

Data were collected in another kindergarten class directly across the hall from the target setting to serve as a social comparison (see Table 3). In other words, these data were used to establish whether the intervention brought the target behavior to a level equal to that of peers who are functioning appropriately (Kazdin, 1982). Five data points were collected during floor time which resulted in a mean response rate of 1.6 talk-outs per minute (range 1.1-2.3). Seven data points were collected during math time which resulted in a mean response rate of 1.5 talk-outs per minute (range 1.2-2.2).

Supplementary Data

Interobserver Agreement

Before assessing the effectiveness of the intervention, interobserver agreement (IOA) was established to ensure that data were collected consistently and reliably throughout the study. Interobserver agreement was calculated using the total count method, dividing the smaller number of responses by the larger number and multiplying by 100 (Kazdin, 1982). IOA was established (and calculated above 80%) prior to the start of the study. To ensure consistent reliability, both researchers coded 6 out of the 12
Table 3

*Comparison of Mean Response Rate of Talking Out*

<table>
<thead>
<tr>
<th></th>
<th>Target Class</th>
<th>Social Comparison Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floor Time Baseline</td>
<td>4.3</td>
<td>1.6</td>
</tr>
<tr>
<td>Floor Time Intervention</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>Floor Time Maintenance</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Math Baseline</td>
<td>3.9</td>
<td>1.5</td>
</tr>
<tr>
<td>Math Intervention</td>
<td>2.2</td>
<td></td>
</tr>
</tbody>
</table>

(50%) baseline data points with IOA ranging between 80% and 97% ($M = 90$). After baseline data were assumed reliable, the second researcher coded 19 out of 56 data points which represented 34% of the treatment and maintenance data in each setting (approximately every third day of data collection). Agreement levels ranged between 68% and 97% ($M = 88$) during the first setting, and between 78% and 95% ($M = 89$) during the second setting. Agreement levels across both settings ranged between 68% and 97% ($M = 89$). The primary researcher and two data based researchers were in contact (i.e., discussed possible clarifications on the target behavior definition) twice because IOA fell below 80% in the target class and three times because of low levels of agreement in the social comparison class. Low levels of agreement were noted on days that talk-outs occurred less frequently.

The social comparison class’s data were also co-coded by two researchers 5 out of 12 data points (approximately every third day), which resulted in 42% of the data
co-coded. Agreement scores ranged between 92% and 94% ($M = 93$) during the first setting, and between 71% and 79% ($M = 74$) during the second setting. Agreement levels across both settings ranged between 71% and 94% ($M = 82$). Tables 4 and 5 display the IOA data.

Table 4

*IOA Results for the Target Class*

<table>
<thead>
<tr>
<th>Setting</th>
<th>Range</th>
<th>Mean ($M$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across Settings Baseline</td>
<td>80-97</td>
<td>90</td>
</tr>
<tr>
<td>Across Settings Treatment/Maintenance</td>
<td>68-97</td>
<td>89</td>
</tr>
<tr>
<td>Floor</td>
<td>68-97</td>
<td>88</td>
</tr>
<tr>
<td>Math</td>
<td>78-95</td>
<td>89</td>
</tr>
</tbody>
</table>

Table 5

*IOA Results for the Social Comparison Class*

<table>
<thead>
<tr>
<th>Setting</th>
<th>Range</th>
<th>Mean ($M$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across Settings</td>
<td>71-94</td>
<td>82</td>
</tr>
<tr>
<td>Floor</td>
<td>92-94</td>
<td>93</td>
</tr>
<tr>
<td>Math</td>
<td>71-79</td>
<td>74</td>
</tr>
</tbody>
</table>
In response to research question 5, the Intensity scores on the SESBI-R and ECBI were analyzed for statistical significance. On average, the Intensity score on the SESBI-R obtained before and after the intervention was below the cutoff score indicating severe child behavior problems at school (Pre-intervention \( n = 23, M = 150.43, SD = 50.88 \); Post-intervention \( n = 23, M = 148, SD = 43.09 \)). Eyberg and Pincus (1999) have suggested scores above 151 as cutoffs to estimate the severity of child behavior problems at school. However, the teacher clearly found the behavior challenging. A paired samples \( t \)-test was performed to compare scores before and after the intervention. There was no statistical difference between Intensity scores on the SESBI-R, \( t(1, 22) = .42, p = .68 \).

A similar result was observed with the SESBI-R Problem scores, which were lower than the suggested clinical cutoff score (Pre-intervention \( n = 23, M = 14.78, SD = 9.89 \); Post-intervention \( n = 23, M = 12.48, SD = 9.69 \)). Scores above 19 represent severe child behavior problems at school (Eyberg & Pincus, 1999). Results from a paired samples \( t \)-test approached significance, indicating a decrease in problem behavior scores over time, \( t(1, 22) = 2.00, p = .06 \).

On average, the Intensity score on the ECBI obtained before and after the intervention was below the cutoff score suggestive of severe child behavior problems at home (Pre-intervention \( n = 12, M = 98.58, SD = 24.21 \); Post-intervention \( n = 12, M = 96.42, SD = 27.36 \)). Eyberg and Pincus (1999) have suggested scores above 131 as cutoffs to estimate the severity of child behavior problems at school. A paired samples \( t \)-test
$t$-test was performed to compare scores before and after the intervention. There was no statistical difference between Intensity scores on the SESBI-R, $t(1, 11) = .40, p = .70$.

ECBI Problem scores were also lower than the suggested clinical cutoff score (Pre-intervention $n = 12, M = 10, SD = 7.32$; Post-intervention $n = 11, M = 5.55, SD = 5.57$) with scores above 15 representing severe child behavior problems at home (Eyberg & Pincus, 1999). Although, a paired samples $t$-test revealed a significant difference between pre and post ECBI Problem scores, $t(1, 10) = 2.82, p = .02$ (see Table 6).

Table 6

*Descriptive Data for SESBI-R and ECBI Behavior Rating Scales*

<table>
<thead>
<tr>
<th></th>
<th>Pre-intervention $M \ (SD)$</th>
<th>Post-intervention $M \ (SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SESBI-R Intensity Score</td>
<td>150.43 (50.88)</td>
<td>148 (43.09)</td>
</tr>
<tr>
<td>SESBI-R Problem Score</td>
<td>14.78 (9.89)</td>
<td>12.48 (9.69)</td>
</tr>
<tr>
<td>ECBI Intensity Score</td>
<td>98.58 (24.21)</td>
<td>96.42 (27.36)</td>
</tr>
<tr>
<td>ECBI Problem Score</td>
<td>10 (7.32)</td>
<td>5.55 (5.57)</td>
</tr>
</tbody>
</table>

*Intervals of Time*

The rate of behavior during each setting was calculated by the researcher on a daily basis by dividing the frequency of behavior by the amount of time in minutes (rounded to the nearest half minute) and then plotted on the graph. At the end of each week, this data was averaged to determine the interval of time for reinforcement (IRT) to be used by the teacher. Table 7 indicates the IRT (what the teacher set her timer at) for
Table 7

*Interresponse Times (IRT) for Floor*

<table>
<thead>
<tr>
<th>Intervention Week</th>
<th>IRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>25 seconds</td>
</tr>
<tr>
<td>Week 2</td>
<td>25 seconds</td>
</tr>
<tr>
<td>Week 3</td>
<td>35 seconds</td>
</tr>
<tr>
<td>Week 4</td>
<td>40 seconds</td>
</tr>
<tr>
<td>Week 5</td>
<td>30 seconds</td>
</tr>
</tbody>
</table>

each week of intervention in the first setting. The intervention was implemented in the second setting for one week using a 40 second IRT.

Conclusion

This study was conducted to answer the above mentioned research questions. In conclusion, although DRO reduced talking out in both settings, it did not reduce talking out to near zero levels. Momentary DRO decreased talking out more effectively than whole DRO when implemented during floor time. Performance feedback coupled with a review of student outcome data (through Microsoft Excel graphs) effectively maintained treatment integrity levels. The acceptability level of the intervention was low as indicated by the BIRS. Finally, results of the perceived intensity of problem behavior by both the teacher and parents, as measured by the SESBI-R and ECBI, did not significantly change as a result of the intervention.
CHAPTER IV
DISCUSSION

Chapter 4 presents a summary of the results when implementing DRO classwide with young children and a discussion of these results in relation to (a) educational links, (b) hypotheses, (c) teacher feasibility, (d) limitations, and (e) implications for practice and future research.

Findings

The DRO intervention was implemented classwide with 24 typically developing kindergarten children across two settings (i.e., floor time and math) in order to decrease levels of talking out. A whole interval application was utilized during the treatment phase and momentary interval during maintenance. Interresponse times were averaged on a weekly basis to enhance teacher feasibility. A reinforcement assessment was conducted to ensure the children were sufficiently motivated to change their behavior.

Talking out decreased during both floor time and math, yet it was not reduced to near zero levels. However, the levels of talking out during math in the target class approached the teacher’s goal (i.e., talking out levels equivalent with the social comparison class). Momentary DRO as implemented during the maintenance phase not only maintained the decrease in talking out, but appeared to have a stronger effect. Treatment integrity (i.e., adherence to the intervention protocol) was high, whereas teacher acceptability of the intervention was low. When comparing subjective data from
before and after the intervention, there was no statistical difference in the teacher’s rating of the students’ intensity of behavior or whether behaviors were viewed as a problem. Similarly, there was no statistical difference in the parents’ rating of the students’ intensity of behavior, but statistical significance was attained in the parents’ rating of problem behavior.

The teacher’s initial goal of reducing the talking out was not achieved and the social comparison class’s talking out proved to remain at a consistently lower rate than the target class. Experimental control was threatened by an observed change in behavior in the second setting when the intervention was introduced in the first setting.

Educational Links

The DRO intervention was chosen because of its link to the empirically supported tenets of Applied Behavior Analysis (ABA) and Positive Behavior Supports (PBS). Research has suggested that the combination of PBS and ABA offers systematic processes to improve socially significant behavior (Cooper et al., 1987). Education has long supported ABA because of its effective modification of behavior and focus on analyzing data, whereas PBS is endorsed by The Individuals with Disabilities Education Act (IDEA, 2004). PBS focuses on positive reinforcement, preventive approaches to problem behavior, implementing research based practice, and changing behavior to a socially important level (Sugai et al., 2000).

Differential Reinforcement of Other Behavior (DRO) is an evidence-based, positive reinforcement intervention that reinforces behavior other than the target behavior and utilizes treatment data to determine time intervals for reinforcement. Although the
current study did not conclude a reduction of the target behavior to near zero levels, the
strength of DRO should not be discounted. Research continues to support that
reinforcement-based behavioral strategies derived from ABA have successfully decreased
challenging behaviors (Lerman, Vorndran, Addison, & Contrucci Kuhn, 2004).

Hypotheses

Upon reviewing the outcomes in light of the existing literature, there appear to be
several possible explanations for the findings of this study. These explanations are
discussed within the context of teacher acceptability data, treatment integrity data,
outcomes of the behavior rating scales, characteristics of the participants, and issues
surrounding reinforcement.

Teacher Acceptability

Witt and Elliott (1985) hypothesized a fundamental link between treatment
acceptability and treatment effectiveness. They argued that if an intervention agent
“accepts” the intervention, it will tend to be more effective. The low teacher acceptability
in this study was one of the most profound pieces of data recorded.

A review of the literature reveals an intervention’s acceptability may be
influenced by psychologist, teacher, and/or child factors (Elliott, 1988). Psychologists
control how the intervention is introduced and how much they are willing to support the
intervention agent. Research reports that the evaluations of treatments (i.e., acceptability)
vary according to what the treatment is called, how it is described by the psychologist
(i.e., jargon), and the psychologists’ level of involvement in the intervention (Elliott).
Teacher variables that influence their preference of treatments are reported in Table 8. Children’s treatment preferences vary considerably across studies, but tend to vary 

Table 8

*Teacher Variables That Influence Treatment Acceptability*

<table>
<thead>
<tr>
<th>Teacher Variable</th>
<th>How it effects acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity of the target behavior</td>
<td>The more severe the behavior, the more acceptable any intervention</td>
</tr>
<tr>
<td>Type of treatment</td>
<td>Positive procedures rated as more acceptable</td>
</tr>
<tr>
<td>Time required to implement the intervention</td>
<td>Teachers prefer more time efficient interventions</td>
</tr>
<tr>
<td>Reported effectiveness of the intervention</td>
<td>Providing treatment effectiveness information leads to more acceptability</td>
</tr>
<tr>
<td>Teacher’s attitude and knowledge of behavioral principles</td>
<td>More knowledge, more acceptability</td>
</tr>
<tr>
<td>Teacher experience</td>
<td>More experience leads to less acceptance of interventions</td>
</tr>
</tbody>
</table>

(Elliott, 1988)

support interventions that emphasize individual student-teacher interactions and group reinforcement; while public reprimands and negative contingencies for an entire group when only student misbehaved were reported unacceptable.

This study attempted to promote high levels of acceptability by manipulating the psychologist variables (i.e., use jargon and involvement), the teacher’s years of experience, the type of treatment and its effectiveness, and the child’s behavior severity. For example, the researcher took precautions to explain the intervention to the teacher using everyday language and to remain involved throughout the study (e.g., physically
present each day of the study, weekly face to face meetings). A first year teacher implemented the intervention which supports the conclusion that teachers with more experience in the field are less likely to find treatments acceptable (Witt & Robbins, 1985). The intervention implemented (DRO) is categorized as a positive reinforcement based intervention and has been proven effective across various behaviors and populations. The class’s severity of behavior was best evidenced using the social comparison class’s data. The only variable that worked against acceptability in this study was the time it took to implement the intervention. The children’s acceptability level was not measured.

Although steps were taken to ensure high levels of acceptability, limitations were noted. For example, the researcher did not actively collaborate with the teacher in choosing the intervention as recommended in the literature (Hughes & DeForest, 1993). This may have led the teacher to feel “left out” of the process and less invested in the study, have less confidence in the effectiveness of the intervention, and feel less motivated to implement the intervention with enthusiasm. At the same time, one cannot overlook the time of year this study took place. Although the teacher expressed an interest in decreasing the talking out in her class, the intervention began in April, which is toward the end of the school year. This time of year may be correlated with an increase in student misbehavior and an increase in teacher burn-out. This may have created a cycle in which an ineffective intervention (one that lacked immediate, positive results) led to less treatment fidelity by the teacher, and the teacher’s low level of enthusiasm may have been part of the reason the intervention was less than effective. In other words, if the
The teacher had implemented the intervention with confidence and enthusiasm, it may have been more effective.

In conclusion, whereas the teacher’s pre-intervention ratings indicated only a slight disagreement with the intervention, her post-intervention ratings signified a strong disagreement with the treatment components. The less than optimal results of this study may reflect the hypothesized fundamental link between treatment acceptability and treatment effectiveness (Witt & Elliott, 1985). In other words, the observed low teacher acceptability may explain the lower than expected results of the intervention.

**Treatment Integrity**

Although this study evaluated treatment integrity using the recommended method of both direct observation and a self-report measure (Goss, Noltemeyer, & Devore, 2007), the results are subject to limitations. High treatment integrity scores indicated that the teacher was comfortable implementing the intervention. Although the teacher successfully completed the nine steps listed on the checklist (resulting in 100% adherence to the protocol), the teacher’s actual adherence may have been inflated. For example, some of the steps (e.g., paired labeled praise with reinforcement/icon, ignored target behavior) involved variable opportunities throughout the study. There may have been 10 opportunities for the teacher to ignore talking out on one particular day, but she may have only ignored it 8 of the 10 times. This was scored a “yes” on the checklist (indicating she had followed the protocol) because she had successfully implemented that particular criteria over half the opportunities presented (more than 50%). More stringent criteria on the checklist may have resulted in a more consistent application of the intervention.
Specifically, each instance of talking out would have been ignored rather than just most of the instances.

In addition, the treatment integrity checklists gave the researcher an insight into the possibility of the teacher being overwhelmed by the study. For example, at the beginning of the intervention, the teacher completed each session’s checklist daily. Toward the end of the study, even with written reminders, the teacher failed to complete them each day, saving some to complete during the post-study meeting.

Relationship between Treatment Acceptability, Integrity and Outcomes

The results of this study support the hypothesis that treatment acceptability and treatment integrity impact intervention outcomes (Witt & Elliott, 1985). Initially in this study, treatment acceptability scores were above average and treatment integrity scores were high. By the conclusion of the study, treatment acceptability fell, treatment integrity scores remained high and the intervention’s outcome data was poor. It is possible that the teacher’s enthusiasm dropped after the realization that her goal was not being met, her perception of the intervention may have changed after implementation, or she may have lost interest because of the difficulty of implementation. Although the teacher continued to implement the intervention as designed, it appeared she did not “buy into” it anymore resulting in a less effective intervention than predicted.

SESBI-R/ECBI Implications

At the onset of the study, the researcher hypothesized that if the intervention produced the expected change in behavior, a corresponding change in behavior would be reflected in the behavior rating scales. Unfortunately, because the intervention was not as
effective as predicted and did not reduce talking out levels to the teacher’s goal or to near
zero levels, no statistical difference was noted in the students’ behavior before and after
the study.

In addition, a comparison between the teacher’s and parents’ behavior rating
scales provided further evidence of the teacher being overwhelmed by her students’
behavior. The teacher’s reported means were close to clinically significant cutoff scores
for intensity of behavior, whereas parents’ reported means were not. In other words, the
teacher rated the students’ behavior as more severe when compared to the parents’
perceptions. This infers that she saw the behavior as more severe than the parents and
therefore, may have been more overwhelmed by it.

Differences in expectations may be a potential explanation for this outcome. The
classroom environment tends to be more structured than the home environment and for
that reason; there may be higher expectations of behavior in the classroom. For example,
certain behaviors (e.g., temper tantrums, teasing, fighting, interrupting) result in different
consequences depending on the setting. Students may exhibit the same behavior in both
settings, but the behavior and its severity impacts the environment differently. It is also
possible that as a first year teacher, she had expectations of her students that were higher
and less practical than those of seasoned teachers who tend to adjust their behavioral
expectations with classroom experience.

Participants

In general, DRO has mainly been implemented with students with disabilities
(e.g., autism, Newman et al., 1997; mental retardation, Deitz et al., 1976; Poling et al.,
1978; and multiple handicaps, Didden et al., 1997). Not only do these students present with different characteristics than the current study’s participants (i.e., typically developing students), but also warranted the intervention to reduce their target behaviors to near zero levels. It was with this past research in mind that the current study established expectations and research questions which may not have been most functional for this population.

As an alternative, it could be argued that consideration of the peer comparison data be used over the near zero expectation as established in past research. Peer comparison, a method to compare the target population with peers who are functioning appropriately (Kazdin, 1982), measured part of the social validity in this study. Social validity would dictate that perhaps the social comparison class provides a more functional goal than near zero levels. Future research may benefit from exploring the use of a social comparison class to establish intervention goals with typically developing populations.

Reinforcement

A last possible explanation for the findings may lie in the issue of reinforcement. It is well documented that reinforcer potency (i.e., its ability to compete with already successful reinforcing inappropriate behavior and variation of reinforcement in order to avoid satiation) are crucial in successfully changing behavior (Mason et al., 1989; Poling & Ryan, 1982; Repp et al., 1983; Repp et al., 1991). Therefore, this study attempted to ensure reinforcer potency by conducting a reinforcement assessment prior to implementing the intervention (Didden et al., 1997) and by providing the class with a
choice of reinforcers at the start of each session of intervention. In spite of this, five issues arose regarding reinforcement.

First, it is possible that the students did not completely comprehend the goal of the reinforcement assessment. For example, results of the reinforcement assessment yielded six options for the students to choose from each session (e.g., high fives, hugs, Goldfish Crackers®, stickers, gummies, and Smarties®). Although hugs and high fives were agreed upon by 95% of the class (84% respectively), they were rarely chosen by the group or individuals during intervention. This infers that the students may not have chosen items that were motivating enough to change their behavior.

Second, the novelty of the reinforcement may have lost its strength. As mentioned earlier, variation of reinforcement is critical in order to avoid satiation with reinforcement (Mazaleski et al., 1993). There were only four options of reinforcement that student typically chose among in this study. Although reinforcers were rotated, four options did not generate a large rotation or selection. To test this hypothesis, the researcher conducted a second reinforcement assessment with items that fell in the 70-79% category during the original assessment (e.g., Skittles®, Coco Puffs®). Ninety-one percent of the class voted for Skittles® (as opposed to the 74% that voted for it initially). Thus, Skittles® were added as a reinforcement choice the last five days of maintenance. A drastic decrease in talking out was observed the first day that Skittles® were offered (0.63 rpm), but similar results were not observed on subsequent days (1.9, 1.7, 4, 2.6). A drastic decrease in the target behavior was also observed the first day of intervention in the first setting which also may be attributed to the novelty of the reinforcement. This data
suggests that the students’ behavior was capable of changing when motivating reinforcement was delivered.

Third, during the intervention phase, the reset interval option was utilized which gives attention to the target behavior and lowers the rate of reinforcement (Vollmer et al., 1993). In other words, the students may have been more motivated by the attention they received from their teacher for their talking out than by the reinforcers they chose. Also, on the several occasions that the levels of talking out remained high, the students’ schedule of reinforcement was low. As a result, the students may have been discouraged by the lack of reinforcement and therefore, lost incentive to perform.

In contrast, during the maintenance phase, a continuous interval option was utilized due to the momentary interval application of DRO. The continuous interval option emphasizes the extinction component of DRO which ignore the target behavior. In this situation, the students’ talking out was completely ignored (i.e., the timer was not reset, thereby not giving attention to the target behavior). Also, because the continuous interval option was paired with a momentary interval application (mDRO), it offered more opportunities for reinforcement because students were rewarded if they were not talking out only at the end of the interval rather than for the entire interval. The students may have found this more rewarding which may explain why the momentary (mDRO) interval application was more effective than the whole (wDRO) interval.

Fourth, research clearly states that delayed reinforcement is less effective than reinforcement delivered immediately following the behavior (Kazdin, 1977). Due to teacher request, beginning on the seventh day of intervention (in the first setting),
students were rewarded on a delayed schedule of reinforcement. Icons were posted instead of the reinforcement being delivered, with students rewarded at the end of the approximately 10 minute lessons. For the same reason, students’ reinforcement was sometimes delayed almost 40 minutes when the intervention was implemented in the second setting (math). Delayed reinforcement for kindergarten aged students may have decreased the potency of the reinforcement and thus, may have reduced the effectiveness of the intervention.

Fifth, it is possible that the reinforcement, and thus, the intervention, was less effective because it was implemented after lunch. The combination of factors that the students commonly chose food as reinforcement (e.g., Gummy Worms®, Goldfish Crackers®) and that the intervention was implemented 30 minutes after lunch, may explain why the students were not as motivated by food to change their behavior. This phenomenon was noted by Corte et al. (1971) who reported the DRO intervention as more effective when implemented under a mild food deprivation condition as compared to a no food deprivation condition.

Teacher Feasibility

The researcher performed several tasks in an attempt to make the intervention more acceptable and feasible for the teacher. For instance, throughout the study, the researcher set up the video camera for taping and was responsible for video taping each of the settings. The researcher also prepared the classroom before the teacher arrived by setting the timer and preparing the reinforcement choices for each setting.
On-the-spot modifications to the intervention were made to allow the lesson to run more smoothly. For example, if the timer went off during a group response (e.g., class counting to 50, singing the days of the week), the students continued the activity and were rewarded afterwards. Also, the teacher expressed concern over the extended amount of time it took to distribute the reinforcement to the entire class and their resulting high levels of distraction once they were rewarded. Initially, the researcher helped distribute the reinforcement, but when the teacher continued to express concern, the intervention was modified to allow the teacher to post icons (i.e., tokens; which represented the reinforcement) rather than directly distributing the reinforcement during the lesson. This resulted in the class being reinforced after the lesson (i.e., delayed reinforcement). The researcher also helped distribute the reinforcement at this time. In situations where the teacher continued to teach (after the intervention time had ceased), the researcher organized the reinforcement so it could be distributed while the students lined up for the bathroom.

Another modification was made with the reinforcement system. After consulting with the teacher, the researcher permitted each student to choose their own reinforcement from the two options rather than the entire class agreeing on one reinforcement. The researcher was in charge of keeping the data, organizing the reinforcement, and distributing it in order to increase teacher feasibility. Although attempts were made to make the intervention more acceptable and feasible, the teacher’s rating, as measured by the BIRS, remained low.
Limitations of the Study

The potential limitations of this study are related to (a) Interobserver Agreement (IOA), (b) treatment integrity components, (c) Functional Behavior Assessment (FBA), (d) experimental control, and (e) generalization.

*Interobserver Agreement (IOA)*

To begin, although acceptable levels of agreement vary (Kazdin, 1982), IOA for the social comparison class during math was low ($M = 74$). Therefore, one cannot be certain that the number of behaviors reported were accurate, and thus, may have provided an inaccurate norm for which to compare the target class’s behavior. Second, although neither data collector was informed of phase changes during the study (to control for observer expectancy; Kazdin, 1982), the intervention phase was apparent due to the teacher’s change in behavior. Thus, observer expectancy may have been activated and revealed in the data.

Third, the IOA data may have been inflated because agreement was calculated using the frequency ratio (agreements/agreements+disagreements x 100; Kazdin, 1982). This method cannot discern if both observers were truly tallying the same occurrences (Alberto & Troutman, 2009). For example, if both data collectors observed for 15 minutes and both data collectors reported 15 behaviors, one cannot be certain that they observed the same 15 behaviors. One data collector may have recorded the 15 behaviors in the first 10 minutes and 0 behaviors in the last 5 minutes, whereas the other data collector may have recorded 1 behavior per minute. To improve on this method, agreement could have been calculated by dividing the observation periods into smaller
time intervals and behaviors could have been recorded per interval. Agreement could have been calculated by finding the agreement between intervals rather than the entire time period (Cooper et al., 2007).

_Treatment Integrity Components_

This study attempted to measure whether performance feedback coupled with a review of student outcome data maintained acceptable levels of treatment integrity. Although high levels of treatment integrity were reported, it was difficult to discern whether it was each of the components on their own (i.e., the self-report, direct observation, performance feedback, or the review of student data), the combination of these components, or neither that generated these results. In other words, it might have been the performance feedback aspect, or seeing a change in the students’ behavior, or that she knew she was being watched that influenced the teacher’s behavior (Witt et al., 1997).

_Functional Behavior Assessment (FBA)_

Researchers have argued that in order for DRO interventions to be most effective, the reinforcer for the target behavior needs to be identified and eliminated in order to deliver reinforcement for the absence of the target behavior (Miltenberger, 2001). Examining the function of behavior allows us to better understand why the behavior is occurring and thus, design the most appropriate way to intervene. A formal FBA was not conducted in this study. Consequently, one cannot be certain that the reinforcer for the talking out was eliminated and that the reinforcement offered to the class appropriately replaced it.
Establishing experimental control is essential in single subject research in order to establish evidence-based practices and control for threats to internal validity (Horner et al., 2005). Specifically, in multiple baseline research designs, experimental control is illustrated at each point the intervention is introduced across settings. Each setting should change only when the independent variable (IV) is introduced. Therefore, a functional relationship only exists if changes are apparent in the specific setting being manipulated by the IV. Settings that have not yet received intervention should remain at baseline levels (i.e., changes in the remaining settings should not be observed; Kazdin, 1982). In this study, a change in behavior was observed in the second setting when the intervention was introduced in the first setting. Therefore, one cannot conclude that experimental control was fully attained, and thus, that the intervention was responsible for the change in student behavior.

In addition, three or more baselines are typically used in order to increase the strength of experimental control in multiple baseline designs (Kazdin, 1982). This study implemented the intervention in only two settings due to time constraints (i.e., end of the school year) and lack of teacher support. Including a third setting may have allowed more solid conclusions that it was indeed the intervention that was responsible for the observed change in behavior.

External Validity/Generalizability

A main concern in single-subject research is the extent to which one study’s results have significance beyond that study (e.g., for different participants, settings,
behaviors, intervention agents; Horner et al., 2005). One cannot be certain that the results obtained in this study will “generalize” to other studies. Therefore, generalization, observing these results with different subjects, across different settings, or behaviors, (Stokes & Baer, 1977) needs to be explored in future studies. The external validity of this study may be enhanced through replication of this study, across different participants, behaviors, and settings (Horner et al.). Also, the intervention agent utilized in this study, specifically a first year teacher, may also limit the generalizability of the findings. Although past research has argued that teachers with more experience tend to shy away from interventions (Witt & Robbins, 1985), this researcher argues that it is possible that a teacher with more experience, and consequently a better ability to multi-task, may have encountered more success with the DRO protocol.

**Implications for Practice**

The results of this study highlight 4 areas that may be of interest for school-based practitioners and educators: the issue of momentary versus whole interval implementation, teacher acceptability, teacher feasibility, and reinforcement.

*Momentary Versus Whole Interval Implementation*

Whereas whole interval application of differential reinforcement (wDRO) has been criticized in the past for being too difficult for intervention agents to implement because of the amount of time spent attending to behavior (Repp et al., 1983), momentary interval application (mDRO) has been reported as more valued due to the reduced amount of time invested (Barton et al., 1986). Earlier research had supported the use of wDRO because of its more dramatic effect on behavior, which makes the results of this
study more interesting to intervention agents and teachers. Fortunately for teachers, the results of this study indicated that talking out decreased more effectively when the teacher used a mDRO application of the intervention (as opposed to a wDRO application). Thus, educators may implement DRO using the momentary application and may observe a successful outcome. More research is needed to further assess this potential.

**Teacher Acceptability**

The results of this study also highlight the crucial component of identifying the appropriate intervention, and in turn, “selling the intervention” to the teacher or other treatment agent. School psychologists understand that interventions matched to the function of a behavior are more likely to be effective, as compared to interventions that are only matched to the type of behavior (Watson & Steege, 2003). It is essential to identify which variables control a behavior in order to design the most effective intervention (Watson & Steege). Once this is accomplished, school psychologists can actively involve the teacher. Research encourages working with teachers in a collaborative role (Hughes & DeForest, 1993) to promote teacher “buy in” and thus, increase acceptability and treatment fidelity (Witt & Elliott, 1985).

**Teacher Feasibility**

One could certainly conclude that DRO is not easily applied classwide in the general education classroom after considering all of the modifications made in this study to accommodate the teacher and attempt to increase teacher feasibility. Practitioners are reminded that successful DRO studies have distributed reinforcement in a classwide
circumstance to randomly selected students rather than the entire class (Daddario et al., 2007). Success has also been routinely reported when pairing DRO with a token economy (Repp et al., 1976). The use of a token economy may increase teacher acceptability and feasibility by allowing the teacher to continue the lesson rather than interrupting to deliver reinforcement.

Reinforcement

The topic of reinforcement continues to be of critical importance to practitioners and researchers alike, since it is well established that reinforcement is the key to effectively teaching and changing behavior. This study attempted to control reinforcement potency and variation as recommended by the literature (Mason et al., 1989; Mazaleski et al., 1993), but alas did not achieve the expected results. The method and outcomes of this study remind practitioners that reinforcement should be powerful enough to compete with reinforcement of inappropriate behavior, it is important to vary available reinforcement (Poling & Ryan, 1982; Repp et al., 1983; Repp et al., 1991), and reinforcement may be more effective when delivered immediately (rather than delaying it).

Future Directions

Future researchers may contribute significantly to the DRO literature by investigating the intervention implemented with different populations or applying it classwide. In addition, future research could examine differences between whole interval (wDRO) and momentary interval (mDRO) application. Future research may also address the effects of teacher acceptability, feasibility, and treatment integrity. Further topics that
could be addressed in future research include calculating interresponse time (IRT) using seconds and conducting functional behavioral assessments (FBA).

Population

Although the results of this study expand the work of previous research with young children (Conyers et al., 2004; Conyers et al., 2003; Daddario et al., 2007), additional applied research with this specific population is warranted. Specifically, research with typically developing children is sparse. In addition, despite the fact that research has documented the effect years of teaching has on implementing interventions (Witt & Robbins, 1985), this line of inquiry needs to be examined more specifically with the DRO protocol.

Classwide

Although this study expanded on the work of previous research implementing behavioral interventions classwide (Conyers et al., 2004; Conyers et al., 2003; Daddario et al., 2007), it did not replicate the results of Daddario et al. as intended. The current findings imply that future research may benefit from defining distinctions in the term “classwide” in reference to how many students are involved. For example, the Daddario et al. study reported success in implementing DRO classwide, but with a group of only 7 students. It is possible that the current study’s classwide application with 24 students was too large. Clarifications on the size of the group with which the intervention is implemented (and deemed effective) may increase replication effects, and thus, increase the validity of DRO.
Further research is also needed to explore the logistics of implementing DRO at the classwide level in the general education setting. For example, future research may explore outcomes associated with decreasing the amount of time spent distributing reinforcement. In addition, future research may explore the value of students choosing their own reinforcement as opposed to the entire class agreeing. Research is also needed to better understand the effectiveness of DRO implemented at the classwide level, as compared to the effectiveness reported when implemented in small groups or individually.

Whole Interval versus Momentary Interval

The results of this study appear to contradict established previous work which reported whole interval (wDRO) as more effective than momentary interval (mDRO) in decreasing behavior (Repp et al., 1983). Therefore, further research is necessary to confirm if one application is truly more effective than the other. It may be possible that one of the applications proves more successful in certain situations, with particular populations, or in specific settings.

Teacher Acceptability and Feasibility

DRO has been criticized for being too difficult for practitioners to implement (Vollmer et al., 1993). The results of this study support the idea that general education teachers may not “accept” this intervention and therefore, may not benefit from its application in their classrooms. Future research can enhance the reputation of the DRO intervention by increasing teacher acceptability and feasibility through modifying protocols (Deitz et al., 1976), utilizing the continuous interval option rather than the reset
option (Gongola, 2008), working with teachers in a collaborative role (Hughes & DeForest, 1993), and offering continued support throughout the intervention. Further research is also warranted to explore the logistics of delivering immediate reinforcement (the most effective use of reinforcement) while simultaneously building in support so the teacher is able to multi-task.

Treatment Integrity

Future research may benefit from examining the components of the treatment integrity package utilized in this study (i.e., teacher checklist, direct observation, performance feedback, and review of student outcome) in order to determine which is responsible for the observed outcome. Additionally, IOA may be collected for treatment integrity data in order to avoid observer drift (Gongola, 2008). In other words, this data will ensure that more than one individual evaluates the intervention agent and that consistent expectations of how the intervention should be implemented (i.e., treatment integrity) are applied.

Interresponse Time (IRT)

Selecting the appropriate interval size is a significant factor in the effectiveness of DRO (Repp et al., 1976; Repp et al., 1991). Past work in DRO has typically calculated the IRT using minutes (e.g., minutes observed/ number of behaviors; Repp et al., 1974). Future research may explore the alternative of using seconds, rather than minutes, which may represent a more accurate representation of time, and hence, provide more opportunities for reinforcement.
Functional Behavioral Assessments (FBA)

Future research may benefit from conducting a formal FBA prior to implementing the intervention in order to identify variables and possible reinforcement for the target behavior. This best practice decision also increases the likelihood that the DRO intervention will be most effective because it was appropriately chosen to match the function of the target behavior.

Conclusion

The results of this study both confirmed and refuted those outcomes reported in past literature. Similar to past research (Conyers et al., 2004), the DRO intervention was not as effective as hypothesized, and consequently, the kindergarten class’s talking out was not reduced to near zero levels. The teacher reported low levels of acceptability of the intervention (Vollmer et al., 1993), but implemented it with integrity. In contrast to past research (Repp et al., 1983), momentary interval (mDRO) was more effective in reducing the levels of talking out as compared to whole interval (wDRO). The reinforcement assessment (Mason et al., 1989) conducted may not have identified the appropriate reinforcement for this group of students.

Although modifications were made to the intervention protocol, the results of this study suggest that DRO may not be feasibly applied classwide in the general education classroom. Additional research appears warranted. Specifically, future research should address: (a) implementing DRO with young, typically developing populations of students; (b) employing intervention agents of varying years of experience; (c) examining the logistics of implementing the intervention classwide; (d) increasing teacher
acceptability and feasibility; (e) evaluating the effectiveness of momentary interval (mDRO) as opposed to whole interval (wDRO); (f) calculating IRT using seconds rather than minutes to increase reinforcement opportunities; (g) differentiating the effective components of the treatment integrity package; and (h) conducting a formal FBA prior to implementation.

Best Practices for the Interim

Considering the outcomes of this study and in anticipation of future research, educators and intervention agents can focus on collaboration, reinforcement, functional behavioral assessments, and momentary interval application to facilitate implementation in the classroom.

First, best practices recommend actively collaborating with the teacher in choosing the intervention (Hughes & DeForest, 1993). This increases teacher acceptability of the intervention, and thus, intends to increase the effectiveness of the intervention. Second, educators have continually acknowledged that reinforcer potency and variation of reinforcement are critical in teaching and changing behaviors (Mason et al., 1989; Poling & Ryan, 1982; Repp et al., 1983; Repp et al., 1991). Educators are encouraged to continue assessing students’ preferences of reinforcement and delivering them in the most effective manner (i.e., immediately following the desired behavior; Kazdin, 1977). Third, best practices promote conducting a functional behavior assessment (FBA) in order to identify the reinforcer for the target behavior, therefore, rendering the DRO intervention most effective (Miltenberger, 2001). Last, educators are encouraged to capitalize on the results reported in the current study which confirm
momentary interval (mDRO) as more effective than whole interval (wDRO) application.
Momentary interval not only reduces the amount of time invested by teachers (Barton et al., 1986), but also affords students more opportunities for reinforcement.
APPENDICES
APPENDIX A

DEFINITIONS OF BEHAVIORAL CONCEPTS AND STRATEGIES
### Definitions of Behavioral Concepts and Strategies

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Reinforcement</td>
<td>Results in an increase in behavior following the presentation of a favorable event.</td>
</tr>
<tr>
<td>Negative Reinforcement</td>
<td>Results in an increase in a behavior following the removal of an aversive event.</td>
</tr>
<tr>
<td>Differential Reinforcement</td>
<td>A procedure that applies the behavioral concepts of reinforcement and extinction to increase desirable target behaviors or decrease undesirable target behaviors.</td>
</tr>
<tr>
<td>Token Economy</td>
<td>A behavior modification program based on operant conditioning which rewards “tokens” for exhibiting desirable behavior which can be exchanged for back-up reinforcers.</td>
</tr>
<tr>
<td>Punishment</td>
<td>Results in a decrease in behavior following the presentation of an aversive event or removal of a favorable event.</td>
</tr>
<tr>
<td>Overcorrection</td>
<td>Reduces inappropriate behavior through an exaggeration of experience. Restitutional overcorrection requires the student to restore or correct the environment that was disturbed to its condition before the disturbance and then improve it beyond its original condition (overcorrecting the environment). Positive-practice overcorrection requires the student to engage in exaggerated practice of appropriate behaviors.</td>
</tr>
<tr>
<td>Response Cost</td>
<td>The removal or loss of a positive reinforcer contingent on inappropriate behavior.</td>
</tr>
</tbody>
</table>

(Alberto & Troutman, 2009; Kazdin, 1977, 2001; Miltenberger, 2001)
APPENDIX B

STAFF AND STUDENT CONSENT FORMS
Implementing a DRO Intervention Classwide: Target Class

Dear Staff,

I want to do research on how a reinforcement based intervention (Differential Reinforcement of Other Behavior; DRO) may reduce misbehavior in your classroom. This study aims to reduce inappropriate behaviors. I want to do this because minimal literature is available on behavioral interventions for young students and I would like to increase your students’ success in the school environment. I would like you to take part in this project. If you decide to do this, you will be asked to help identify reinforcers (rewards) for your students, identify problem behaviors that are interfering with your instruction, and classroom activities throughout the day that are impacted by misbehavior. Your responsibilities will be to turn on/off the video camera, offer your class a reinforcer to work towards in each activity, and comply with the intervention protocol. You will be learning and implementing the intervention. The DRO intervention, a positive reinforcement-based protocol, is designed to reduce target behaviors while simultaneously increasing time on-task and compliance. It aims to reward students before they misbehave. Each time your students (as a whole class) engage in appropriate classroom behavior, they will earn a reinforcer and will receive verbal praise. If any of them are disruptive, no consequence will be given, and another opportunity to behave appropriately will be given. The entire study will last until June 2009.

Although I do not anticipate it happening, the biggest risk of this project is that your students could become dependent on the DRO intervention in order to behave appropriately. To minimize this, the DRO protocol naturally increases the expectations for good behavior. In other words, as your students’ behavior improves, they will be expected to behave appropriately for longer and longer periods of time in order to receive a reward. Eventually, your class may comply with expectations for an entire day or until they don’t need a reward anymore.

Confidentiality will be maintained to the limits of the law. Confidentiality may not be maintained if your students indicate that they may do harm to themselves or may do/have done harm to others. Information obtained in this study will be treated with complete confidentiality. It will be shared among members of the research team to help develop recommendations for your class upon completion of the study. With your permission after the conclusion of the study, the results will be shared with officials from your school. If the results of this study are published or presented to the officials from your school or at scientific meetings, no information will be presented that could possibly identify you or your students.

If your students take part in this project, I anticipate an increase in on task and compliant behavior in the school setting. The outcome of this project may help other educators use positive reinforcement options when designing interventions. Taking part in
this project is entirely up to you, and no one will hold it against you if you decide not to do it. If you do take part, you may stop at any time.

If you want to know more about this research project, please call me at 440.842.0436 or my advisor, Dr. Richard Cowan, at 330.672.2294. The project has been approved by Kent State University. If you have questions about Kent State University's rules for research, please call Dr. John West, Acting Vice President and Dean, Division of Research and Graduate Studies (Tel. 330.672.2704).

You will get a copy of this consent form.

Sincerely,

Rosie Daddario, M.Ed. School Psychology Doctoral Candidate
Implementing a DRO Intervention

I agree to take part in this project. I know what I will have to do and that I can stop at any time.

_______________________________________________   _______________________
Signature       Date

Video Consent Form

Video will only be used to help count how many times the class misbehaves and to see if you are following the DRO protocol (our data collection). The video will be uploaded to a private blog on the internet and only those who are given rights (me, 2 data collectors, and my advisor) will be able to view it. The videos will be destroyed and the blog deleted when the study is over.

I agree to video taping at _____ NHA________________________________________ from February 2009- June 2009.

_______________________________________________   _______________________
Signature       Date

I have been told that I have the right to see the video tapes before they are used. I have decided that I:

____ want to see the tapes   ______ do not want to see the tapes

Sign now below if you do not want to see the tapes. If you want to see the tapes, you will be asked to sign after seeing them.

Rosie and other researchers approved by Kent State University may/may not (circle one) use the tapes made of my class. The original tapes or copies may be used for:

____ this research project       ____ teacher education       ____ presentation at professional meetings

_______________________________________________   _______________________
Signature       Date
Dear Parents/Guardians,

I want to do research in your child’s classroom on how a reinforcement based intervention (DRO) may reduce misbehavior. I want to do this because only a few people have studied behavioral interventions for young students and I would like to increase your child’s success in school. I would like you to let your child take part in this project. If you decide to do this, your child will choose something that motivates them (also called a reinforcer) to work towards to encourage them to follow the intervention. The DRO intervention focuses on rewarding your child before they misbehave. It is designed to decrease misbehaviors while also increasing the time your child spends doing school activities and listening to the teacher. Each time your child acts appropriately, they will earn a reinforcer (such as a sticker or pretzel) and will receive verbal praise (such as “nice job keeping your hands to yourself”). If they are disruptive, they will not be punished, and another opportunity to earn a reinforcer will be given. The entire study will last until the end of the year (June 2009).

Although I do not expect it to happen, the biggest risk of this project is that your child could become dependent on the DRO intervention in order to behave appropriately. Expectations for good behavior will increase as your child’s class becomes better behaved. They will be expected to behave for longer and longer periods of time in order to receive a reward. Eventually, your child may comply with expectations for an entire day or until they don’t need a reward anymore.

Confidentiality will be maintained to the limits of the law. Confidentiality may not be maintained if your child indicates that s/he may do harm to themselves or may do/have done harm to others. Information obtained in this study will be treated with complete confidentiality. It will only be shared among members of the research team to help develop recommendations for your child’s class. With your permission after the conclusion of the study, the results will be shared with officials from your school. If the results of this study are published or presented to the officials from your school or at scientific meetings, no information will be presented that could possibly identify you or your child.

If your child takes part in this project, I anticipate an increase in good behavior during school. The results of this project may encourage other teachers use positive reinforcement options (rewarding good behavior) to decrease misbehavior. Taking part in this project is entirely up to you, and no one will hold it against your child if you decide not to do it. If your child does take part, s/he may stop at any time.

If you want to know more about this research project, please call me at 440.842.0436 or my advisor, Dr. Richard Cowan, at 330.672.2294. The project has been approved by Kent State University. If you have questions about Kent State University's
rules for research, please call Dr. John West, Acting Vice President and Dean, Division of Research and Graduate Studies (Tel. 330.672.2704).

You will get a copy of this consent form.

Sincerely,

Rosie Daddario, M.Ed. School Psychology Doctoral Candidate
Implementing a DRO Intervention

I agree to let my child take part in this project. I know what s/he will have to do and that s/he can stop at any time.

_______________________________________________ _______________________
Signature       Date

**Video Consent Form**

Video will only be used to help count how many times the class misbehaves and to see if the teacher is rewarding the students correctly (our data collection). The video will be uploaded to a private blog on the internet and only those who are given rights (me, 2 data collectors, and my advisor) will be able to view it. The videos will be destroyed and the blog deleted when the study is over.

I agree to video taping at _____NHA________________________________________
from February 2009- June 2009.

_______________________________________________ _______________________
Signature       Date

I have been told that I have the right to see the video tapes before they are used. I have decided that I:

_____ want to see the tapes       _____ do not want to see the tapes

Sign now below if you do not want to see the tapes. If you want to see the tapes, you will be asked to sign after seeing them.

Rosie and other researchers approved by Kent State University **may/may not** (circle one) use the tapes made of my child. The original tapes or copies may be used for:

_____ this research project       _____ teacher education       _____ presentation at professional meetings

_______________________________________________ _______________________
Signature       Date
Implementing a DRO Intervention Classwide: Social Comparison Class

Dear Staff,

I want to do research on how a reinforcement based intervention (Differential Reinforcement of Other Behavior; DRO) may reduce misbehavior in the classroom. This study aims to reduce inappropriate behaviors. I want to do this because minimal literature is available on behavioral interventions for young students and I would like to increase students’ success in the school environment. I would like you to take part in this project. If you decide to do this, your class will be used as a comparison to the target kindergarten class. Your responsibilities will be to turn on/off the video camera and then return it to Ms. Celmar. I ask that you change nothing you do. In other words, conduct class as you typically do. The entire study will last until June 2009.

Confidentiality will be maintained to the limits of the law. Confidentiality may not be maintained if your child indicates that s/he may do harm to themselves or may do/have done harm to others. Information obtained in this study will be treated with complete confidentiality. It will only be shared among members of the research team. With your permission after the conclusion of the study, the results will be shared with officials from your school. If the results of this study are published or presented to the officials from your school or at scientific meetings, no information will be presented that could possibly identify you or your child.

If you want to know more about this research project, please call me at 440.842.0436 or my advisor, Dr. Richard Cowan, at 330.672.2294. The project has been approved by Kent State University. If you have questions about Kent State University's rules for research, please call Dr. John West, Acting Vice President and Dean, Division of Research and Graduate Studies (Tel. 330.672.2704).

You will get a copy of this consent form.

Sincerely,

Rosie Daddario, M.Ed. School Psychology Doctoral Candidate
Implementing a DRO Intervention
I agree to take part in this project. I know what I will have to do and that I can stop at any time.

_______________________________________________ ________________________
Signature       Date

Video Consent Form
Video will only be used to help count how many times the class misbehaves (our data collection). The video will be uploaded to a private blog on the Internet and only those who are given rights (me, 2 data collectors, and my advisor) will be able to view it. The videos will be destroyed and the blog deleted when the study is over.

I agree to video taping at ___NHA_________________________________________
from February 2009- June 2009.

_______________________________________________ ________________________
Signature       Date

I have been told that I have the right to see the video tapes before they are used. I have decided that I:
_____ want to see the tapes   _____ do not want to see the tapes

Sign now below if you do not want to see the tapes. If you want to see the tapes, you will be asked to sign after seeing them.
Rosie and other researchers approved by Kent State University may/may not (circle one) use the tapes made of my child. The original tapes or copies may be used for:

_____ this research project   _____ teacher education   _____ presentation at professional meetings

_______________________________________________ ________________________
Signature       Date
Implementing a DRO Intervention Classwide: Social Comparison Class

Dear Parents/Guardians,

I want to do research in the other kindergarten class at your child’s school on how a reinforcement based intervention (DRO) may reduce misbehavior. I want to do this because only a few people have studied behavioral interventions for young students and I would like to increase students’ success in school. **Although your child’s class will not be receiving this intervention, I would like you to let your child take part in this project. If you decide to do this, your child’s class will be videotaped once a week for 45 minutes in order to count their misbehaviors. We will use your child’s class as a comparison to the other kindergarten class. The entire study will last until the end of the year (June 2009). There is no risk to your child.**

Confidentiality will be maintained to the limits of the law. Confidentiality may not be maintained if your child indicates that s/he may do harm to themselves or may do/have done harm to others. Information obtained in this study will be treated with complete confidentiality. It will only be shared among members of the research team. With your permission after the conclusion of the study, the results will be shared with officials from your school. If the results of this study are published or presented to the officials from your school or at scientific meetings, no information will be presented that could possibly identify you or your child.

If you want to know more about this research project, please call me at 440.842.0436 or my advisor, Dr. Richard Cowan, at 330.672.2294. The project has been approved by Kent State University. If you have questions about Kent State University's rules for research, please call Dr. John West, Acting Vice President and Dean, Division of Research and Graduate Studies (Tel. 330.672.2704).

You will get a copy of this consent form.

Sincerely,

Rosie Daddario, M.Ed. School Psychology Doctoral Candidate
Implementing a DRO Intervention

I agree to let my child take part in this project. I know what s/he will have to do and that s/he can stop at any time.

_______________________________________________
_________________________
Signature       Date

Video Consent Form

Video will only be used to help count how many times the class misbehaves (our data collection). The video will be uploaded to a private blog on the internet and only those who are given rights (me, 2 data collectors, and my advisor) will be able to view it. The videos will be destroyed and the blog deleted when the study is over.

I agree to video taping at ___NHA___________________________________________ from February 2009- June 2009.

_______________________________________________
_________________________
Signature       Date

I have been told that I have the right to see the video tapes before they are used. I have decided that I:

_____ want to see the tapes   _____ do not want to see the tapes

Sign now below if you do not want to see the tapes. If you want to see the tapes, you will be asked to sign after seeing them.

Rosie and other researchers approved by Kent State University may/may not (circle one) use the tapes made of my child. The original tapes or copies may be used for:

_____ this research project       _____ teacher education       _____ presentation at professional meetings

_______________________________________________
_________________________
Signature       Date
Differential Reinforcement of Other Behavior Applied Classwide with Young Children

Target Class
Assent Script

1. Hi, my name is Rosie, and I am trying to learn more about how to get kids to act the right way in school.

2. I would like you to behave during school. Your teacher will give you prizes if you act the right way. I will be videotaping you so others can see how good you are acting.

3. Do you want to do this?

4. Do you have any questions before we start?

5. If you want to stop at any time just tell me.
Differential Reinforcement of Other Behavior Applied Classwide with Young Children

Social Comparison Class
Assent Script

1. Hi, my name is Rosie, and I am trying to learn more about how to get kids to act the right way in school.

2. I would like to video tape you during school. I don’t want you to do anything different than usual.

3. Do you want to do this?

4. Do you have any questions before we start?

5. If you want to stop at any time just tell me.

A witness statement can be added if the extra protection provided by it is desired.
Data Collection Sheet

<table>
<thead>
<tr>
<th>Date</th>
<th>Class Activity (floor, math)</th>
<th>Length of Activity (in minutes &amp; seconds)</th>
<th>Tallies of behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FITS DEFINITION</th>
<th>DOES NOT FIT DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loud enough to understand</td>
<td>Mumbling, whispering</td>
</tr>
<tr>
<td>Talking or making sounds w/o hand raised &amp; called on</td>
<td>Hand raised &amp; called on</td>
</tr>
<tr>
<td>Off topic conversations</td>
<td>When teacher allows group response</td>
</tr>
<tr>
<td>Repeating another student’s answer</td>
<td>Answers to teacher’s rhetorical questions (“Should I set my timer?”)</td>
</tr>
<tr>
<td>Self-talk above conversation level</td>
<td>Quiet self-talk</td>
</tr>
<tr>
<td></td>
<td>Talking/noises while transitioning to seat</td>
</tr>
</tbody>
</table>
APPENDIX D

INTERVENTION TREATMENT INTEGRITY CHECKLIST
Intervention Phase

Treatment Integrity Checklist

Y  N

θ  θ  1. Choice of reinforcer offered at beginning of activity

θ  θ  1. Beginning of the activity introduced by saying “Let’s begin (activity).”

θ  θ  3. Timer initially set

θ  θ  4. Reinforcement delivered/icon posted at end of interval (if target behavior was not observed)

θ  θ  5. Labeled praise coupled with reinforcement/icon

θ  θ  6. Timer reset after reinforcement delivered/icon posted

θ  θ  7. Inappropriate behavior ignored

θ  θ  8. Reset timer if inappropriate behavior was observed

θ  θ  9. End of the activity announced by saying “(Activity) is finished.”

Date: ______________________

Completed by: ______________________

Setting: floor math
APPENDIX E

MAINTENANCE TREATMENT INTEGRITY CHECKLIST
Maintenance Phase

Treatment Integrity Checklist

Y   N

Θ  Θ  1. Choice of reinforcer offered at beginning of activity

Θ  Θ  1. Beginning of the activity introduced by saying “Let’s begin (activity).”

Θ  Θ  3. Timer initially set

Θ  Θ  4. Reinforcement delivered at end of interval (if target behavior was not observed

   at the time of the buzzer)

Θ  Θ  5. Labeled praise coupled with reinforcement

Θ  Θ  6. Timer reset after reinforcement delivered

Θ  Θ  7. Inappropriate behavior ignored

Θ  Θ  9. End of the activity announced by saying “(Activity) is finished.”

Date: ______________________

Completed by: ______________________

Setting:  floor  math
APPENDIX F

BEHAVIORAL INTERVENTION RATING SCALE (BIRS)
Behavior Intervention Rating Scale- Teacher Version

Please evaluate the intervention by circling the number which best describes your agreement or disagreement with each statement. Use 1 indicating you strongly disagree with the statement, 3 indicating a neutral response, and 6 indicating you strongly agree with the statement. Please answer each question.

1. This would be an acceptable intervention for the child’s problem behavior. 1 2 3 4 5 6
2. Most teachers would find this intervention appropriate for behavior problems. 1 2 3 4 5 6
3. The intervention should prove effective in changing the child’s problem behavior. 1 2 3 4 5 6
4. I would suggest the use of this intervention to other teachers. 1 2 3 4 5 6
5. The child’s behavior problem is severe enough to warrant the use of this intervention. 1 2 3 4 5 6
6. Most teachers would find this intervention suitable for behavior problems. 1 2 3 4 5 6
7. I would be willing to use this in the classroom setting. 1 2 3 4 5 6
8. The intervention would not result in negative side-effects for the child. 1 2 3 4 5 6
9. The intervention would be appropriate intervention for a variety of children. 1 2 3 4 5 6
10. The intervention is consistent with those I have used in classroom settings. 1 2 3 4 5 6
11. The intervention was a fair way to handle the child’s problem. 1 2 3 4 5 6
12. The intervention was reasonable for the behavior problem described. 1 2 3 4 5 6
13. I like the procedures used in this intervention. 1 2 3 4 5 6
14. This intervention was a good way to handle this child’s problem. 1 2 3 4 5 6
15. Overall, the intervention was beneficial for the child. 1 2 3 4 5 6
16. The intervention quickly improved the child’s behavior. 1 2 3 4 5 6
17. The intervention would produce a lasting improvement in the child’s behavior. 1 2 3 4 5 6

18. The intervention would improve the child’s behavior to the point that it would not noticeably deviate from other classmate’s behavior. 1 2 3 4 5 6

19. Soon after using the intervention, the teacher noticed a positive change in behavior problems. 1 2 3 4 5 6

20. The child’s behavior will remain at an improved level even after the intervention is discontinued. 1 2 3 4 5 6

21. Using the intervention should not only improve the child’s behavior in the classroom, but also in other settings (e.g., other classrooms, home). 1 2 3 4 5 6

22. When comparing this child with a peer before and after use of the intervention, the child’s and the peer’s behavior would be more alike after using the intervention. 1 2 3 4 5 6

23. The intervention should produce enough improvement in the child's behavior so the behavior no longer is a problem in the classroom. 1 2 3 4 5 6

24. Other behaviors related to the problem also are likely to improve. 1 2 3 4 5 6

BISP Project: Revised 8-21-06; Adapted from BIRS (Von Brock & Elliot, 1987)
APPENDIX G

GOAL ATTAINMENT SCALING (GAS)
Goal Attainment Scale Worksheet

Teacher Report

Target Behavior:

Setting:

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>+2</td>
<td>____________</td>
</tr>
<tr>
<td></td>
<td>____________</td>
</tr>
<tr>
<td>+1</td>
<td>____________</td>
</tr>
<tr>
<td></td>
<td>____________</td>
</tr>
<tr>
<td>0</td>
<td>____________</td>
</tr>
<tr>
<td></td>
<td>____________</td>
</tr>
<tr>
<td>-1</td>
<td>____________</td>
</tr>
<tr>
<td></td>
<td>____________</td>
</tr>
<tr>
<td>-2</td>
<td>____________</td>
</tr>
<tr>
<td></td>
<td>____________</td>
</tr>
</tbody>
</table>

Adapted from GAS (Sladeczek et al., 2001)
REFERENCES
References


