EFFECTS OF A CLASS-WIDE POSITIVE PEER REPORTING INTERVENTION ON MIDDLE SCHOOL STUDENT BEHAVIOR

A dissertation presented by

Ruth Kathryn Chaffee

Submitted to

The Department of Applied Psychology

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the field of

School Psychology

Northeastern University

Boston, Massachusetts

August, 2018
TABLE OF CONTENTS

Title Page ............................................................................................................................... i

Table of Contents .................................................................................................................. ii

Acknowledgements............................................................................................................... vi

List of Tables ........................................................................................................................ vii

List of Figures ....................................................................................................................... viii

List of Appendices ............................................................................................................... ix

CHAPTER I: LITERATURE REVIEW

Abstract .................................................................................................................................. 1

Problem Behavior in Schools ................................................................................................. 2

Approaches to Addressing Problem Behavior in Schools ....................................................... 3

Punitive Approaches to Addressing Student Behavior ......................................................... 4

Positive Approaches to Supporting Student Behavior .......................................................... 6

Positive Behavioral Interventions and Supports ................................................................. 7

Tier 1 Supports ..................................................................................................................... 8

Classroom Management Research ...................................................................................... 10

Group Contingencies .......................................................................................................... 12

Independent Group Contingencies ..................................................................................... 13

Interdependent Group Contingencies ................................................................................ 14

Good Behavior Game ......................................................................................................... 15

Dependent Group Contingencies ......................................................................................... 16

Limitations of Extant Group Contingency Research ......................................................... 17

Tootling ............................................................................................................................. 20
CHAPTER II: EFFECTS OF A CLASS-WIDE POSITIVE PEER REPORTING INTERVENTION ON MIDDLE SCHOOL STUDENT BEHAVIOR

Abstract ...........................................................................................................47

Introduction ....................................................................................................48

Tootling .........................................................................................................50

Purpose of Study ...........................................................................................54

Method ..........................................................................................................56

Participants and Setting ...............................................................................56

Materials .......................................................................................................57

Dependent Variables .....................................................................................57

Class-wide Behavior .....................................................................................57

Social Validity ...............................................................................................59

Teacher-reported Usability ..........................................................................60

Student-reported Usability ...........................................................................60

Study Design and Procedures .....................................................................61

Pre-Baseline ..................................................................................................61

Baseline ........................................................................................................61

Teacher Training ..........................................................................................62
Intervention ................................................................. 62
Withdrawal ................................................................. 64
Reimplementation ...................................................... 64
Maintenance ............................................................... 64
Interobserver Agreement ............................................. 65
Treatment Integrity ...................................................... 65
Data Analysis .................................................................. 67
NAP .............................................................................. 68
Tau-U ............................................................................ 68
Results ........................................................................... 69
Classroom A ................................................................... 69
Classroom B ................................................................... 70
Social Validity ............................................................... 72
Teacher-reported Usability ............................................ 72
Student-reported Usability ............................................ 72
Discussion ....................................................................... 72
Effect of Tootling on Class-wide Behavior .................... 73
Strategies for Generalization of Behavioral Change ....... 75
Social Validity ............................................................... 76
Limitations and Directions for Future Research .............. 78
Implications for Practice .............................................. 80
References ..................................................................... 81
TABLES ......................................................................... 90
FIGURES ................................................................. 95
APPENDICES ............................................................ 96
ACKNOWLEDGEMENTS

First, I would like to express my deepest gratitude to my advisor, mentor, and dissertation committee chair, Dr. Amy Briesch. For the past five years, Amy has provided me endless wisdom, encouragement, patience, and feedback, and has inspired me to be a better school psychologist, researcher, and person. I seek to emulate Amy’s true passion and joy for her work, and I look forward to future collaborations.

I would also like to thank my dissertation committee members, Dr. Robert Volpe, Dr. Laura Dudley, and Dr. Austin Johnson. Each member of my committee provided crucial feedback and guidance that was invaluable to this project. I sincerely appreciate their time, energy, and commitment. I also wish to thank the entire faculty at Northeastern University for their teaching and mentorship.

This study would not have been possible without the financial support of the Society for the Study of School Psychology Dissertation Grant Award and the committed participation of the middle school teachers. I am very grateful for the hours of observations and hard work of my research assistants, Kristin Nissen and Taicha Cornelio. Thank you also to the members of the Center for Research in School-Based Prevention (CRiSP) for their inspiration and support.

Finally, I am eternally grateful for the love and support of my family. Thank you to my parents for nurturing my inquisitive mind, prioritizing my education, and teaching me to care for individuals and communities. Thank you to my wife, Lauren Graber, for her love, encouragement, support, and patience. Our shared passion for social justice and improving the trajectories of lives inspires me every day. Thank you to my children, sweet Sydney and Benny, for keeping me grounded with hugs, laughter, and art projects.
LIST OF TABLES

Table 1. Demographic Data for Each Participating Classroom ..................................................90
Table 2. Overall Percentages of Observer-Scored Treatment Integrity ......................................91
Table 3. Effect Size Calculations ..............................................................................................92
Table 4. URP-IR Mean (SD) by Subscale and Classroom .........................................................93
Table 5. CURP Mean (SD) by Subscale and Classroom ..............................................................94
LIST OF FIGURES

Figure 1. Effects of Tootling on Middle School Classrooms ......................................................95
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>Tootling Intervention—Demographic Form</td>
<td>96</td>
</tr>
<tr>
<td>Appendix B</td>
<td>Quiz on Operational Definitions of Dependent Variables</td>
<td>98</td>
</tr>
<tr>
<td>Appendix C</td>
<td>Tootling Behavioral Observation Form</td>
<td>100</td>
</tr>
<tr>
<td>Appendix D</td>
<td>Usage Rating Profile-Intervention (Revised)</td>
<td>101</td>
</tr>
<tr>
<td>Appendix E</td>
<td>Children's Usage Rating Profile</td>
<td>103</td>
</tr>
<tr>
<td>Appendix F</td>
<td>Tootling Intervention Scripts</td>
<td>105</td>
</tr>
<tr>
<td>Appendix G</td>
<td>Procedural Integrity Checklist—Observer Form—Teacher Training</td>
<td>108</td>
</tr>
<tr>
<td>Appendix H</td>
<td>Procedural Integrity Checklist—Observer Form—Student Training</td>
<td>109</td>
</tr>
<tr>
<td>Appendix I</td>
<td>Procedural Integrity Checklist—Observer Form—Intervention</td>
<td>110</td>
</tr>
<tr>
<td>Appendix J</td>
<td>Procedural Integrity Checklist—Observer Form—Withdrawal</td>
<td>111</td>
</tr>
<tr>
<td>Appendix K</td>
<td>Procedural Integrity Checklist—Observer Form—Maintenance</td>
<td>112</td>
</tr>
<tr>
<td>Appendix L</td>
<td>Procedural Integrity Checklist—Teacher Form—Intervention</td>
<td>113</td>
</tr>
<tr>
<td>Appendix M</td>
<td>Procedural Integrity Checklist—Teacher Form—Withdrawal</td>
<td>114</td>
</tr>
<tr>
<td>Appendix N</td>
<td>Procedural Integrity Checklist—Teacher Form—Maintenance</td>
<td>115</td>
</tr>
</tbody>
</table>
CHAPTER I

Abstract

Off-task and disruptive classroom behavior are significant barriers to student academic achievement. Within multi-tiered systems of support, class-wide interventions, such as group contingencies, have emerged as effective and feasible methods of addressing and preventing disruptive behavior within an entire class. However, there is limited research of class-wide interventions implemented at the secondary level, many of the existing configurations of interventions are negatively focused, and specific intervention configurations need additional research. Tootling has emerged as a promising, class-wide intervention which encourages positive behavior through positive peer reporting and an interdependent group contingency.
LITERATURE REVIEW

Problem Behavior in Schools

Disruptive behaviors are the most common reasons for school-based disciplinary referrals (Sterling-Turner, Robinson, & Wilczynski, 2001), out of school suspensions, and expulsions (Rausch & Skiba, 2004). In a recent Schools and Staffing Survey, a national integrated survey of public and private schools, 38% of teachers reported that student misbehavior interfered with their teaching (National Center for Education Statistics, 2012). By the middle and high school level, behavioral problems often become more severe. Of secondary school teachers surveyed by Public Agenda, 70% reported that disruptive behavior is a serious concern and over 75% indicated that their teaching would be more effective if time spent managing disruptive behavior was reduced (Public Agenda, 2004).

Childhood disruptive behavior and interpersonal difficulties are significant barriers to both academic success and achievement (Atkins, Hoagwood, Kutash, & Seidman, 2010; Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Levitt, Saka, Hunter Romanelli, & Hoagwood, 2007). Disruptive behavior has been associated with more significant behavioral and mental health needs and future educational achievement, delinquency, substance abuse, and adult criminal behavior (Greer-Chase, Rhodes, & Kellam, 2002; Gresham, 1985; Parker & Asher, 1987; Wentzel, 1991). Students getting out of their seats, making noises, talking out of turn, arguing, and failing to follow classroom rules and demands can deter other students’ ability to focus on academic material, disrupt the sense of classroom safety, and increase teachers’ levels of stress (Malecki & Elliot, 2002; Walker, Ramsey, & Gresham, 2003/2004). Additionally, students exhibiting disruptive behaviors demand significant teacher attention, which can interfere
with the instruction of all students and the recognition of positive student behavior (Luiselli, Putnam, & Sunderland, 2002; Sterling-Turner et al., 2001).

In addition to disrupting the learning environment, pervasive disruptive behaviors can impact the social-emotional and behavioral health of all students in the classroom. Exposure to negative peer behavior and a lack of structure in the classroom have been shown to have a contagion effect and increase the risk of later aggression and disruptive behavior for all students in the class (Kellam, Ling, Merisca, Brown, & Ialongo, 1998). Furthermore, when disruptive behavior is left unaddressed, there are greater long-term risks both for those students exhibiting the behaviors and those exposed to it. Research indicates that conduct-related issues may have a critical period of intervention in which the longer children go without access to effective intervention, the more intractable and expensive the issues are to treat (Bradley, Doolittle, & Bartolotta, 2008; Gresham, 1991; Loeber & Farrington, 1998). Given the extensive impacts for both affected individuals and peers in the classroom, approaches are needed to address disruptive behavior in school classrooms.

**Approaches to Addressing Problem Behavior in Schools**

Emmer and Stogh (2001) define classroom management as “actions taken by the teacher to establish order, engage students, or elicit their cooperation” (p. 103). In the past, teachers and schools primarily utilized punishment-based classroom management techniques, such as threats, paddling, removal from class, suspension, or expulsion, to reduce disruptive student behavior (Lyman, 2006). Though the threat or execution of the punishment reduces disruptive behavior for some students, for others it can also cause increased behavioral difficulty and covert misbehaviors, such as vandalism (B. F. Skinner, 1968; C. H. Skinner et al., 2000). Furthermore, simply discouraging negative behaviors does not produce enduring positive behavioral change.
Over the past few decades, research has shown that explicit teaching and reinforcement of appropriate replacement behaviors produces long-term improvements in behavior (LeGray, Dufrene, Sterling-Turner, Olmi, & Bellone, 2010; Volmer & Iwata, 1992; Volmer, Roane, Ringdahl, & Marcus, 1999). Thus, proactive and positive classroom management approaches, such as within Positive Behavioral Intervention and Supports (PBIS), in which appropriate behaviors are preventatively taught and encouraged for all students, are more appropriate, effective, and permanent method of addressing disruptive classroom behavior (Sugai & Horner, 2002).

**Punitive Approaches to Addressing Student Behavior**

Historically, schools have typically emphasized punitive techniques in order to address student problem behavior. Punitive school discipline practices were established and reinforced both by laws and the Supreme Court. For example, corporal punishment, physical punishment intended to inflict pain (e.g., striking a student on the buttocks with a wooden paddle, rapping knuckles with a ruler), was widely used in public schools in the past. As of 1971, only two states—New Jersey and Massachusetts—had banned the practice (Lyman, 2006). As recently as 1977, the Supreme Court upheld the constitutional right to disciplinary corporal punishment in public schools in *Ingraham v. Wright*. Corporal punishment is still common in some areas of the South; more than 167,000 public school students received corporal punishment in the 2011-2012 school year (Anderson, 2015). In the 1990s, fear of crime and a public misperception of school violence ushered in zero-tolerance student discipline policies (i.e., specific, harsh punishment for rule infractions regardless of the circumstances), and the incorporation of police and police procedures in schools including searches of lockers and students, staff and student identification cards, and prosecuting juvenile offenders as adults (Hyman & Perone, 1998). In 1994, the federal
Gun-Free Schools Act instituted a mandatory year-long expulsion for any student bringing a weapon to school. Since then, state-based codes of conduct have commonly included zero-tolerance policies for guns, violence, and drugs and mandatory penalties for violations, rather than proactive or positive reinforcement based approaches (American Psychological Association Zero Tolerance Task Force, 2008; Barton, Coley, & Wenglinsky, 1998). These policies have focused on mandated- and often severe- consequences for specific actions, regardless of the gravity of the behavior, under the assumption that removing students committing these disruptive actions will deter other students and improve the overall school climate (Public Agenda, 2004).

Within this context, most systems for managing student behavior in schools and other educational settings have tended to emphasize monitoring and surveillance, rules, and reactive punishment strategies (Skinner, Cashwell, & Skinner, 2000; Sugai & Horner, 2002). In the classroom, teachers often establish punishment-based programs with specific consequences for violating established class rules (e.g., student who swears lose recess for a week). Such punishment-based classroom management strategies, where the rules and consequences are pre-established, are aimed at preventing disruptive or anti-social behavior. To some extent, they have been shown to reduce some future occurrence of antisocial behaviors (Sulzer-Azaroff & Mayer, 1986). However, punishment-based systems also incite students, particularly students with established behavioral difficulties, to escalate behavior or develop maladaptive strategies to avoid punishment (e.g., covert behaviors, vandalism; Martin & Pear, 1992; Mayer, 1995; Mayer & Butterworth, 1979; Mayer, Butterworth, Nafpaktitis, & Sulzer-Azaroff, 1983; B. F. Skinner, 1968; C. H. Skinner et al., 2000). Furthermore, although these behavior management techniques are designed to facilitate academic instruction, increased rates of suspension and expulsion have been linked to less satisfactory school climate, increased amount of time spent on disciplinary
matters, and lower school-wide academic achievement, even when controlling for socioeconomic status (Davis & Jordan, 1994; Raffaele-Mendez, Knoff, & Ferron, 2002; Scott & Barrett, 2004; Skiba & Rausch, 2006). Finally, these punitive strategies typically do not include any teaching or reinforcement of appropriate replacement behaviors, which have been found to be effective long-term strategies for addressing disruptive behavior (LeGray et al., 2010; Volmer & Iwata, 1992; Volmer et al., 1999). Given these concerns with aversive and reactive strategies, focus has increasingly shifted toward positive approaches to manage disruptive behavior.

**Positive Approaches to Supporting Student Behavior**

Over the past few decades, research, policy, and practice in school-based discipline have been slowly shifting away from use of punishment-based strategies to a more positive, proactive focus in regards to student social emotional and behavioral health. In *The Teaching of Technology* (1968), Skinner lamented teachers’ use of aversive control techniques in the classrooms including scolding, loss of privileges, and detention. In 1991, and again in 2001, the American Academy of Pediatrics publicly argued against the use of corporal punishment in schools, citing that the practice adversely affects student self-image, academic achievement, and disruptive behavior (Committee on School Health, 1991, 2000). The shift toward more positive classroom management was also spurred by changes in federal special education identification policy with the Individuals with Disabilities Education Act (IDEA) of 1997, the No Child Left Behind Act (NCLB) of 2001, the Individuals with Disabilities Education (IDEA) Act of 2004, and most recently, Every Student Succeeds Act (ESSA) of 2015 (Sugai & Horner, 2002). These policies established Response to Intervention (RtI), now more commonly incorporated into multi-tiered systems of support (MTSS), as the practice of providing high-quality and evidence-based instruction and intervention aligned with student need, frequent progress monitoring, and
data-based decision making (Batsche et al., 2005). MTSS utilizes a public health prevention framework to prevent behavior and academic problems for the majority of students and intervene effectively for students with more challenging behaviors or persistent academic difficulties (Sugai & Horner, 2006). It includes three tiers of supports: (a) primary prevention (Tier 1), including school-wide and class-wide systems and supports directed towards all students such as effective classroom management and academic support; (b) targeted interventions (Tier 2), such as small-group administered social skills, self-management for those who require increased adult attention and monitoring; and (c) intensive individual supports (Tier 3), often administered by special educators, counselors, school psychologists, and behavior interventionists to the small group of students (5%) who are unresponsive to Tier 1 and 2 supports. In response to the federal legislation, school systems began adopting a MTSS to prevent and address academic, mental, emotional, and behavior problems in students. Positive Behavioral Interventions and Supports (PBIS) is one of the more popular multi-tiered frameworks for supporting student behavior; however, adoption of MTSS for social emotional and behavioral concerns has been much slower than for academic concerns (Saeki et al., 2011; SpectrumK12, 2011; Walker, Severson, Feil, Stiller, & Golly, 1998).

**Positive behavioral intervention and supports.** One articulation of the shift from punitive to positive approaches to addressing student behavior has been through PBIS, which employs a continuum of universal and individualized strategies to support academic and social-emotional outcomes by preventing problem behavior (Sugai & Horner, 2002; Sugai et al., 2000). The central elements of PBIS include a focus on clearly defined expectations, evidence-based interventions, and praise and attention for appropriate behaviors (Lewis & Sugai, 1999). PBIS acknowledges that children come to school with previous experiences (e.g., trauma or harsh
behavior management practices) that may contribute to problem behavior. Furthermore, the model eschews aversive school disciplinary practices which have been shown to exacerbate rates of problem behavior, and instead articulates a proactive prevention and early intervention approach (Lewis & Sugai, 1999).

At the primary prevention level, PBIS seeks to prevent the development of problem behavior by explicitly teaching appropriate or desired behaviors and establishing a school and classroom environment that facilitates learning and reduces the chance of problem behaviors. Strategies at this tier include carefully designing the physical layout of the classroom, explicitly teaching and practicing classroom routines (e.g., turning in work, arrival, dismissal), providing reminders of expected appropriate behaviors, and using specific praise when students exhibit appropriate behavior (Simonsen et al., 2015). Secondary prevention strategies are focused on reducing the current cases of problem behavior by providing targeted group-based strategies for small-groups that are at-risk or are already exhibiting problem behavior (Lewis & Sugai, 1999). Finally, tertiary prevention involves individualized behavior supports designed to reduce the intensity and/or severity of well-established problem behaviors (Sugai & Horner, 2006; Sugai et al., 2000). Though PBIS is an evidence-based approach, less than 10% of schools nationwide are actually implementing its tiered behavioral supports (www.pbis.org). Based on the research, it would seem the remaining 90% of schools would benefit from feasible, evidence-based interventions in order to provide tiered behavioral supports.

**Tier 1 supports.** Tier 1 or primary prevention supports, the classroom management strategies a teacher uses, are a critical level of intervention. Not only are these supports provided to all students, but employing effective interventions at this level may prevent more significant mental health problems and reduce long-term costs and impacts associated with the identification
and treatment of more impairing and serious disabilities (National Research Council & Institute of Medicine, 2009). Within PBIS and MTSS, tier 1 behavioral supports rely heavily on class-wide interventions (Riley-Tillman & Burns, 2009). Class-wide behavioral interventions are effective teaching and classroom management strategies implemented by teachers with all students in a classroom to improve classroom behavior (e.g., reduce disruptive behavior, increase student engagement). Class-wide interventions provide opportunities for strong student relationships, modeling of appropriate behaviors, and feedback for all students, and teachers find them more acceptable as opposed to individual interventions (Fantuzzo & Atkins, 1992; State, Harrison, Kern, & Lewis, 2017). Furthermore, effective class-wide interventions have been shown to reduce disruptive behavior and prevent later aggression, delinquency, substance use, and special education service use by at-risk students (Bradshaw, Zmuda, Kellam, & Ialongo, 2009; Chaffee, Briesch, Johnson, & Volpe, 2017; Kellam et al., 1998; Kellam, Rebok, Ialongo, & Mayer, 1994).

However, despite the fact that classroom management and disruptive behavior has long been recognized as a critical issue in schools, teachers have consistently reported that they do not feel adequately trained to manage these issues (Anderson & Kincaid, 2005; Emmer & Stough, 2001; Greer-Chase et al., 2002; Public Agenda, 2004; The New Teacher Project, 2013). In a recent review of teacher preparation programs, The National Council of Teacher Quality found that although a majority of programs do address classroom management, the content is perfunctory rather than an attempt to instill research-informed practice (Greenberg, Putman, & Walsh, 2014). Consistent with this lack of training, research has found that the interventions used in schools to address disruptive behavior are primarily based on personal experience, rather than empirical evidence (Bramlett, Murphy, Johnson, Wallingsford, & Hall, 2002). Thus, providing
teachers with additional knowledge and training in regards to effective classroom management strategies may allow for early intervention and the possible prevention of future behavioral health problems, as well as the management skills to increase instructional time. It is critical that evidence-based class-wide behavioral interventions be identified, disseminated, and implemented by teachers struggling to manage disruptive behaviors. The use of class-wide behavioral interventions may alleviate the disruptive behavior as well prevent any compounding of symptoms or the necessitation of more intense service delivery outside of the school setting. Although increased attention has been paid to these goals in recent years, early classroom management research was more descriptive in nature and continued research is needed.

**Classroom Management Research**

There has been an expanding body of research on classroom practices and behavior management strategies since the 1960s. Early research on classroom management focused on correlational or descriptive studies of effective teachers (Emmer & Stough, 2001; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). However, over the past fifty years the use of sound experimental design to demonstrate effectiveness of interventions, the establishment and expansion of applied behavioral analysis, and the use of appropriate analytical procedures have led to the identification and validation of several types of class-wide interventions that are effective at managing student behavior in classrooms (e.g., reducing disruptive behavior, increasing engagement; Baer, Wolf, & Risley, 1968).

Most recently, class-wide intervention techniques to address disruptive classroom behavior have been identified via meta-analyses and research syntheses of a particular intervention (e.g., token economies, Maggin, Chafouleas, Goddard, & Johnson, 2011; Good Behavior Game, Bowman-Perrott, Burke, Zaini, Zhang, & Vannest, 2016) as well as larger
syntheses of multiple class-wide interventions. In recent years, several meta-analyses and research syntheses have focused on describing the evidence for particular class-wide intervention strategies. For example, Bowman-Perrott and colleagues (2016) conducted a meta-analysis of the Good Behavior Game (GBG; Barrish, Saunders, & Wolf, 1969). Results found that the GBG had a large effect on reducing disruptive behaviors and increasing appropriate behavior, with significantly greater effects found for students with or at risk for emotional/behavioral disabilities and those exhibiting increased rates of disruptive and off-task behaviors. Additionally, several research syntheses on group contingency interventions, both more broadly and specifically analyzing token economies, have found these interventions have a large effect at reducing disruptive behavior (Maggin et al., 2011; Maggin, Johnson, Chafouleas, Ruberto, & Berggren, 2012; Maggin, Pustejovsky, & Johnson, 2017). Maggin et al. (2017) found that overall, group contingencies meet What Works Clearinghouse (WWC) criteria for evidence-based practice (Kratochwill et al., 2010).

Though these research syntheses and meta-analyses have highlighted effective interventions, there are several limitations to note. First, some of these analyses did not restrict included studies to class-wide applications of the intervention (e.g., Maggin et al., 2011; Maggin et al., 2012), or to the general education classroom context. As these contexts can vary widely (e.g., two targeted students in a special education classroom with three teachers vs. all students in a general education classroom with one teacher), it is difficult to generalize the results to a general education class-wide context. Furthermore, these research syntheses and meta-analyses have noted poor study design (e.g., not reporting IOA), and poor demographic reporting regarding participating students in included studies (Bowman-Perrott et al., 2016; Maggin et al., 2011; Maggin et al., 2012; Maggin et al., 2017). Finally, as these studies have employed
different inclusion criteria, procedures, and effect size metrics, it is difficult for teachers to compare across interventions. Thus, given these limitations, research syntheses of individual interventions are difficult for teachers to use to determine an appropriate research-based intervention to use in their general education classroom context.

Several studies have synthesized the research on multiple class-wide behavioral interventions (Oliver, Wehby, & Reschly, 2011; Stage & Quiroz, 1997); however to date only one has systematically identified the evidence base for class-wide behavioral interventions conducted specifically in the general education classroom context (Chaffee et al., 2017). Chaffee and colleagues (2017) identified 29 single-case design (SCD) studies for inclusion in their meta-analysis of class-wide interventions for supporting student behavior in the general education classroom. Although a number of different interventions and configurations were identified (e.g., Good Behavior Game, GBG; Barrish et al., 1969, color wheel, self-management, peer tutoring), it is of note that 25 of the 29 identified studies included some form of group contingencies. Results indicate that class-wide behaviorally oriented interventions are effective overall at both reducing disruptive behavior and increasing appropriate behavior (Tau-U=.93; Hedges’ g=2.04). Though further moderator analyses regarding specific configurations of interventions were limited by the number of studies, the Good Behavior Game, interdependent contingencies, and independent group contingencies (token economies) were all found to be similarly effective class-wide behavioral interventions.

**Group Contingencies**

The majority of studies identified by Chaffee et al. (2017) utilized group contingencies as an intervention component (e.g., Christ & Christ, 2006; Crouch, Gresham, & Wright, 1985; Greenwood, Hops, Delquadri, & Guild, 1974; Kelshaw-Levering, Sterling-Turner, & Skinner,
2000; Ling, Hawkins, & Weber, 2011; McKissick, Hawkins, Lentz, Hailley, & McGuire, 2010; Robichaux & Gresham, 2014). With group contingencies, group membership parameters determine whether positive reinforcement (e.g., rewards) or punishment procedures are implemented. By creating linked goals, students can learn to work together toward their goal and group members can serve as reminders to focus on the targeted behaviors (Davies & Witte, 2000). Advantages of group contingencies over individual contingencies include reduced teacher time for group monitoring and reinforcing behaviors, and the potential for more students to receive reinforcement (Kamps et al., 2011). The group contingency is among the most researched classroom management procedure with strong evidence supporting its efficacy in improving student behavior (Maggin et al., 2012; Maggin et al., 2017; Stage & Quiroz, 1997). Studies implementing class-wide group contingencies have also included self-management procedures (e.g., Hoff & Ervin, 2013), video modeling (e.g., Battaglia, Radley, & Ness, 2015) and multiple contingencies (e.g., Solomon & Tyne, 1979). The three types of group contingency procedures include interdependent, dependent, and independent (Litow & Pumroy, 1975).

**Independent group contingencies.** In independent group contingencies, such as token economies, individual students earn rewards (e.g., tokens, points) contingent upon a desired behavior (e.g., each student who behaves appropriately for five minutes of class time earns one point). These tokens can later be exchanged for back-up reinforcers such as desired items, free time, or privileges (Kazdin, 1977). Within the meta-analysis, Chaffee et al. (2017) identified six studies that utilized independent group contingencies, either alone or in combination with other intervention components (Coogan, Kehle, Bray, & Chafouleas, 2007; Crouch et al., 1985; Mandlebaum, Russell, Krouse, & Gonter, 1983; McLaughlin & Malaby, 1972; Ringer, 1973; Solomon & Tyne, 1979; Thorne & Kamps, 2008). Only three studies within the meta-analysis
involved only a token economy (McLaughlin & Malaby, 1972; Ringer, 1973; Solomon & Tyne, 1979). As one example of a token economy, McLaughlin and Malaby (1972) examined the effect of a reward points system for quiet behavior on inappropriate verbalizations in a fifth and sixth grade classroom. Each student earned a point for every five minutes of appropriate class behavior. Points could be exchanged for classroom privileges such as playing a sport, or taking out special playground equipment. Inappropriate verbalizations decreased significantly. Although past analyses have found a lack of support for token economies, largely due to a failure of studies to meet study design standards (Maggin et al., 2011), Chaffee and colleagues (2017) found that token economies had an overall large to very large effect on classroom behavior (Tau-U = .90).

Interdependent group contingencies. In interdependent group contingencies, reinforcement is contingent on a combined group performance against a criterion (e.g., if a class as a whole has no more than four occurrences of talking out in a period, then all students earn 10 extra minutes of recess). With an interdependent group contingency in place, students may be both personally motivated and feel peer pressure to earn the reinforcer. Chaffee et al. (2017) identified ten studies that utilized interdependent group contingencies, either alone or in combination with other intervention components or contingencies (Christ & Christ, 2006; Coogan et al., 2007; Crouch et al., 1985; Greenwood et al., 1974; Kelshaw-Levering et al., 2000; Ling et al., 2011; Mandlebaum et al., 1983; McKissick et al., 2010; McLaughlin & Malaby, 1972; Ringer, 1973; Robichaux & Gresham, 2014; Solomon & Tyne, 1979; Thorne & Kamps, 2008). These authors found that the seven studies that implemented a basic interdependent group contingency in the absence of other components had an overall large to very large effect on classroom behavior (Tau-U = .95). As one example, Ling and colleagues (2011) implemented an
interdependent group contingency in a first grade classroom during the morning meeting and activities on the rug. After reminding students of the behavioral expectations, the teacher would evaluate the class’ behavior at three separate times during the session. If all students were behaving appropriately, the class would earn a smiley face for the board. If the class earned all three possible smiley faces, the students would receive a beanie baby stuffed animal on their desks immediately following morning rug activities. Both the target student and peers showed improvements in engagement and off-task behavior when the intervention was in place.

**Good Behavior Game.** The most popular example of an interdependent group contingency is the Good Behavior Game (GBG; Barrish et al., 1969), which has been recommended by the Surgeon General as a Promising Program for the prevention of youth violence (U.S. Department of Health and Human Services, 2001). Chaffee et al. (2017) found that the eight studies that utilized the GBG had an overall large to very large effect on classroom behavior (Tau-U = 1.00) (Barrish et al., 1969; Donaldson, Vollmer, Krous, Downs, & Berard, 2011; Donaldson, Wiskow, & Soto, 2015; Kleinman & Saigh, 2011; McGoey, Schneider, Rezzetano, Prodan, & Tankersley, 2010; Nolan, Filter, & Houlihan, 2013; Tanol, Johnson, McComas, & Cote, 2010). In its original form, Barrish and colleagues (1969) split the class into two teams, with each team collectively earning marks for each rule violation (e.g., talking out, being out-of-seat). The team with the fewest marks or both teams if they received more than the pre-established criterion, earned a reward (e.g., 30 minutes of free time, lining up first for lunch). The GBG leverages peer influence to assist in managing student behavior pressure; rather than gaining social attention (e.g., laughs, smiles) from peers for disruptive behavior, in the context of the game peers are more likely to withhold attention or provide disapproval (Tingstrom, Sterling-Turner, & Wilczynski, 2006).
The original GBG and various modifications have proven highly effective at reducing disruptive behavior in the classroom setting, as well as increasing appropriate social behaviors for wide age ranges of students (Barrish et al., 1969; Donaldson et al., 2011; Donaldson et al., 2015; Kleinman & Saigh, 2011; McGoe et al., 2010; Nolan et al., 2013; Tanol et al., 2010). A recent meta-analysis of SCD GBG studies found that the intervention is more effective at reducing off-task behavior than at increasing on-task behavior and that behavioral improvements are more significant for students with more severe emotional and behavioral concerns (Bowman-Perrot et al., 2016). The GBG has also been shown to have long-term prevention effects, reducing the likelihood of later aggression, shy behaviors, and substance usage for students exposed to the intervention in first grade (Dolan et al., 1993; Kellam et al., 1994). Given the evidence supporting the immediate and preventative impact of the GBG, it has even been deemed a “universal behavioral vaccine,” which should be implemented as a public health measure (Embry, 2002, p. 274).

**Dependent group contingencies.** In dependent group contingencies, the rewards or outcomes for the group are determined based on whether select members of a group meet the criterion (e.g., if Johnny has no more than four occurrences of talking out in a period, then all students earn 10 extra minutes of recess). Dependent group contingencies have been critiqued as resulting in negative levels of peer pressure or even retaliation for the targeted students (Davis & Blankenship, 1996). Chaffee and colleagues (2017) did not identify any studies that solely employed dependent group contingencies; however two studies employed dependent contingencies as one component of a larger intervention package (Battaglia et al., 2015; Coogan et al., 2007). Battaglia, Radley, and Ness (2015) implemented On-Task in a Box (Jenson & Sprick, 2014), a manualized, practice-ready intervention including video modeling of appropriate
behavior, student self-monitoring of on-task behavior, and a dependent group contingency. Following a 20-minute session of implementing the video modeling and self-monitoring procedures, the teacher would randomly collect five students’ self-monitoring forms. If the average on-task behavior of those five students was above the predetermined criterion, all students in the class earned a reward.

**Limitations of extant group contingency research.** Although there is much research to support the use of group contingency interventions to support student behavior at the class-wide level, there are nonetheless limitations that are worth acknowledging.

First, recent analyses of class-wide group contingency interventions have noted limited research conducted at the secondary level (Chaffee et al., 2017; Maggin et al., 2017). Of the 29 studies included in the Chaffee et al. (2017) meta-analysis, only four were conducted at the secondary level (Christ & Christ, 2006; Kleinman & Saigh, 2011; McDonnell, Mathot-Buckner, Thorson, & Fister, 2001; Mitchem, Young, West, & Benyo, 2001). Similarly, the recent comprehensive research synthesis of group contingencies across settings (Maggin et al., 2017) identified only three studies published in peer reviewed journals that were conducted in secondary school general education classes (Christ & Christ, 2006; Dart, Radley, Battaglia, & Dadakhodjaeva, 2016; Mitchell, Tingstrom, Dufrene, Ford, & Sterling, 2015). Thus, very limited research on class-wide interventions has been conducted at the secondary level. This is a significant limitation because the teaching and learning context in middle and high school differs significantly from elementary school (e.g., a rotating schedule of different teachers is used at the secondary level as compared to students spending the entire school day with one teacher at the elementary level). Furthermore, middle and high school students are in a unique developmental phase of early adolescence in which students are trying to increase autonomy and identity
development, which may lead to rejection of class-based interventions as an assertion of student independence or because the interventions are deemed too juvenile (Akos, 2005). Finally, effective interventions for secondary students are particularly needed as students with social-emotional and behavioral difficulties present with these symptoms later in their academic career and with more severe symptoms than students who are identified earlier (Walker, Nishioka, Zeller, Severson, & Feil, 2000). Thus, interventions successfully conducted at the primary level cannot necessarily be generalized to the secondary level, and additional research is needed.

Second, nearly half of the interventions identified in the Chaffee et al. (2017) meta-analysis employed punishment-based procedures (e.g., response cost in the GBG; Barrish et al., 1969; Donaldson et al., 2011; Donaldson et al., 2015; Kleinman & Saigh, 2011; McGoey et al., 2010; Nolan et al., 2013; Tanol et al., 2010) or focused attention on negative student behaviors (Coogan et al., 2007; Crouch et al., 1985; Kelshaw-Levering et al., 2000; Mandlebaum et al., 1983; McKissick et al., 2010; Robichaux & Gresham, 2014; Solomon & Tyne, 1979). Punishment-based procedures may exacerbate the behavioral difficulties for at-risk students (Martin & Pear, 1992; Mayer, 1995; Mayer & Butterworth, 1979; Mayer et al., 1983; B. F. Skinner, 1968; C. H. Skinner et al., 2000). Furthermore, more enduring behavior change has been shown to result from teaching and reinforcing appropriate behaviors (LeGray et al., 2010; Volmer & Iwata, 1992; Volmer et al., 1999). Thus, although the interventions have evidence supporting their efficacy, many of the procedures are not focused on increasing appropriate replacement behaviors. In recent years, the field has begun to examine at positively focused interventions, and modifications of existing interventions, to encourage appropriate behavior and increase alignment with PBIS.

One example of modifications being made to a class-wide intervention to better align
with the PBIS framework can be seen in the GBG. For example, Tanol and colleagues (2010) compared the traditional response-cost version of the GBG, in which rule-violations receive teacher attention and a negative mark, and a positive modification of the GBG, wherein students earn points and the teacher provides praise for rule-following behavior. Both methods were effective at reducing disruptive student behavior, and teachers reported preferring the positively focused version. Next, Wright and McCurdy (2012) implemented both the GBG and an adaptation of the GBG, the Caught Being Good Game (CCGG; Wolf, Hanley, King, Lachowicz, & Giles, 1970) in a kindergarten and fourth grade classroom. In the positive adaptation, the class is split into two teams, with each team earning points contingent on the entire team engaging in expected classroom behavior during a variable 20-minute interval. Both the original GBG and the positive variation were effective at decreasing disruptive and increasing on-task student behavior. Teachers and students rated both forms of the GBG as acceptable. Finally, in a replication, Wahl and colleagues (2016) found that both the GBG and the positive modification of the GBG and CCGG are equally effective at improving student behavior. Though this body of research exemplifies the shift towards more positively focused interventions, the research in this area remains limited.

Third, there are several interventions identified by Chaffee and colleagues (2017), which have undergone limited evaluation. The meta-analysis sought to provide conclusive information about the most effective class-wide interventions; however, moderator analyses were not conducted if fewer than three studies utilized the same intervention. Thus, although the effect sizes found in the individual reviewed studies were strong overall, over a third of the studies (n=11), and some of the identified interventions (e.g., Tootling, Color Wheel, On-Task in a Box,
Peer Tutoring), were excluded from further comparisons across interventions. Thus, it is difficult to draw any definitive conclusions regarding the effectiveness of those interventions.

At present, recent meta-analyses have confirmed the efficacy of group contingency interventions at improving student classroom behavior (Chaffee et al., 2017; Maggin et al., 2017). However, a number of limitations exist in the current research base. There has been very limited research conducted at the secondary level, many identified group contingency class-wide interventions employ punishment-based procedures or focus on negative student behaviors, and many of the promising intervention configurations have a limited research base. However, one of the class-wide interventions identified by Chaffee and colleagues (2017), a form of positive peer reporting called Tootling (Skinner, Skinner, & Cashwell, 1998), shows particular promise. Although only two studies in Chaffee et al. (2017) examined the effectiveness of Tootling, it involves a group contingency at its foundation and is aligned with the positive shift in classroom management.

**Tootling**

Rooted in the behavior analytic principles of peer-based reinforcement of appropriate behavior, extinction of inappropriate behavior, and interdependent group contingencies, Tootling involves students reporting their peers’ prosocial behavior (e.g., opening doors, giving positive verbal comments, helping peers, sharing materials; Skinner et al., 1998). Both classroom-based and school-wide prevention have employed peer-to-peer written praise to successfully improve student social competence, academic achievement, violence, aggression, physical health, and vandalism (e.g., Cabello & Terrell, 1994; Embry, Flannery, Vazsonyi, Powell, & Atha, 1996; Mayer et al., 1983; Mayer et al., 1993). Tootling builds upon the strength of peer-to-peer written praise with public feedback and the peer pressure of the interdependent group contingency.
Aligned with the PBIS principles of positive reinforcement of appropriate behavior and prevention, Tootling was initially implemented as an intervention both to teach and to increase the frequency of peer prosocial behavior. Under Bandura’s Social Learning Theory, a potential consequence of having peers monitor each other’s behavior is that students can learn behaviors through observation and imitation, especially when students observe positive peer behavior being reinforced (Bandura, 1965; Bandura, Ross, & Ross, 1963). Social skills that are learned in a natural social environment are also more likely to be generalized and maintained than those learned through more structure planned instruction (Gresham, 1995). Thus, initial implementations of the intervention (Cashwell, Skinner, & Smith, 2001; Skinner et al., 2000) measured the change in the number of student reports of peer prosocial behavior. More recent Tootling studies (Cihak, Kirk, & Boon, 2009; Lambert, Tingstrom, Sterling, Dufrene, & Lynne, 2015; Lum, Tingstrom, Dufrene, Radley, & Lynne, 2017; McHugh, Tingstrom, Radley, Barry, & Walker, 2016) have assessed the impact of the intervention on disruptive and academically engaged behavior.

Skinner and colleagues’ (1998) Tootling procedures involved having elementary students privately report “tootles,” specific instances of peer prosocial behavior, on index cards. “Tootles” were collected daily by the teacher, who read aloud several examples and gave public praise and positive reinforcement for the identified prosocial behavior. Both student reporting behavior and overall student classroom behavior were mediated by an interdependent group contingency. In an interdependent group contingency, where students are affected by peer performance, individuals may indirectly or directly encourage positive, reward-earning behaviors and discourage inappropriate behavior that may negatively impact access to the group rewards (Skinner, Skinner, & Sterling-Turner, 2002). Tootling has also included a visual cue (e.g., thermometer)
displaying cumulative class tootles and progress toward the reward goal. The public visual progress feedback may stimulate peers and teachers to provide additional encouragement or social praise for prosocial behaviors (Seymour & Stokes, 1976; Van Houten, 1984).

As Tootling is largely peer mediated, it does not place significant demands on teachers and has been considered an acceptable intervention by implementing teachers (Lambert et al., 2015; Lum et al., 2017). Further, the theory and components of Tootling are congruent with the key criteria of classroom-based positive behavior support and the MTSS outlined by IDEA (2004) and the recent educational legislation, ESSA (2015). The primary treatment goal is prevention-based with a focus on encouraging positive interactions and behaviors for all students (Murphy & Zlomke, 2014).

**Prior Tootling studies.** To date, six studies have been conducted with the Tootling intervention, all in the southeastern United States (Cashwell et al., 2001; Cihak et al., 2009; Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016; Skinner et al., 2000). Overall, the intervention was implemented in eight elementary classes and three high school classes with class sizes ranging from 17 to 29 students. The four studies reporting student disability status all included general education classes with students with disabilities (e.g., Autism, Other Health Impairment, and Specific Learning Disability). The intervention has demonstrated success in decreasing disruptive student behavior and increasing both reports of prosocial behavior and academically engaged behavior.

Whereas Skinner and colleagues (2000) conducted the initial published study of the Tootling intervention in a fourth grade general education classroom in the rural South, Cashwell and colleagues (2001) conducted a replication study in a second grade classroom. Both studies employed similar procedures and an ABAB withdrawal design. The experimenters
initially trained the students in how to identify and report instances of peer prosocial behavior on index cards. During baseline, experimenters collected tootles from students at the end of each day. The following day, experimenters reported to students the total number of tootles from the prior day and provided students positive and negative examples of tootles from their index cards. Following a stable baseline, an interdependent group contingency and publicly posted feedback were introduced. A ladder poster recording the cumulative total number of tootles was placed on the wall. The class earned a 30-minute recess session for each 100 tootles. The criterion for earning the reward increased to 150 tootles after the initial reward. Across the two studies, results indicated that direct instruction could be used to teach students to report peers’ prosocial behaviors and that the number of tootles was maintained by the group contingency. Some informal evidence for the social validity of the procedures was provided by the teacher’s continued use of the tootling program following the removal of the formal study. Neither study included measurements to assess if the tootling intervention had other impacts such as increasing students’ awareness of prosocial behavior, decreasing anti-social behavior, increasing peer relationships, or improving students’ overall perceptions of school.

Unlike prior studies of Tootling (Cashwell et al., 2001; Skinner et al., 2000), which measured the number of tootles, Cihak and colleagues (2009) extended the evidence on Tootling by measuring the effect of the intervention on the frequency of student disruptive behavior. The intervention was implemented in an ABAB withdrawal design in a third grade class of 19 students, of which four students were identified as having mild disabilities (e.g., specific learning disability, ADHD). A reward criterion of 75 tootles was set, which is notably lower than criterions of 100 and 150 tootles utilized in prior studies (Cashwell et al., 2001; Skinner et al., 2000). Event recording was used to record the total number of disruptive behavior occurrences
performed by the whole class through the school day (8:00am to 2:00pm). Although no effect size was calculated, visual analysis indicated that both the trend and level of disruptive behavior improved with the implementation of Tootling. During the baseline phase, students engaged in a mean of 23.2 (range 19-28) disruptive behaviors daily as compared to the initial Tootling phase in which the mean fell to 8.4 (range 3-15) disruptive behaviors. When Tootling was withdrawn, disruptive behavior increased ($M = 16$, range 19-28), but reduced again with the reimplementation of Tootling ($M = 3.5$, range 0-12). These reductions in disruptive behavior were noted for both students with and without disabilities. Cihak et al. (2009) also extended the evidence on Tootling by including a formal measure of social validity for teachers, the Intervention Rating Profile (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985). Ratings on the IRP-15 indicated that teachers found the intervention acceptable.

This study is limited in that it was focused solely on the impact of Tootling on disruptive behavior, and did not include any assessment of the intervention’s impact on positive or appropriate student behaviors. Furthermore, student treatment acceptability was not assessed and thus there is no indication of how the students perceived the intervention. Finally, like the prior two Tootling studies, no effect size calculations were made to quantify the change in behavior during the intervention. The lack of an effect size statistic prevents comparison of the Tootling intervention to studies of other interventions.

Whereas Cihak et al. (2009) focused exclusively on the effects of Tootling on disruptive behavior, Lambert and colleagues (2015) implemented the Tootling intervention with a fourth and fifth grade classroom to determine the impact on the instances of both disruptive behaviors (e.g., out of seat without permission, inappropriate vocalizations, and unrelated motor movements) and appropriate behavior (e.g., being actively involved or attending to the classroom
task) exhibited by students. The study employed an ABAB withdrawal design with a multiple baseline element across two fourth and fifth grade classrooms. Students were primarily African American and Caucasian and one student in each class received Special Education services under the disability category of Specific Learning Disability. Both disruptive and appropriate behavior were measured for 20 minutes using a 10-second momentary time sampling procedure in which the observed student rotated each interval until all students were observed. In both classrooms, results indicated that the intervention had a moderate to strong effect in reducing disruptive behavior (Nonoverlap of All Pairs (NAP) = .88-1.00) and a strong effect in improving appropriate behavior (NAP= .90-1.00) at the class-wide level. Similar to the limitation of prior Tootling studies, Lambert et al. (2015) did not include any measure of student treatment acceptability.

McHugh and colleagues (2016) further extended the research base on Tootling by assessing the impact of Tootling on disruptive behavior and engagement for both the entire class and individual target students. Additionally, this study was the first and only Tootling study to date to assess student treatment acceptability. Participants were students in three second and third grade classes, as well as one general education target student per class identified by the teacher as exhibiting more disruptive behavior than his/her classmates. The study employed an ABAB withdrawal design with a multiple baseline element across two classrooms and also assessed the impact of a more frequent schedule of reinforcement with the interdependent group contingency. Whereas prior peer reviewed studies of Tootling used classroom criteria of 65-100 tootles (Cashwell et al., 2001; Cihak et al., 2009; Lambert et al., 2015), McHugh et al. (2016) set the criterion for reinforcement at 25-30 tootles, making the reinforcement attainable on a daily basis. Across the three classrooms, results indicated that the intervention had a moderate to strong
effect in reducing disruptive behavior (NAP = .92-1.00) and a strong effect in improving academically engaged behavior (NAP = .98-1.00) at the class-wide level. Results were similar for the target students which showed a moderate to strong effect in reducing disruptive behavior (NAP = .92-1.00) and increasing academically engaged behavior (NAP = .92-1.00). Results did not indicate that a daily goal for tootles, and thus the potential to access reinforcement and rewards more immediately, improved the effect size of the intervention beyond what was seen in the study by Lambert and colleagues (2015). Despite promising results, this study did provide any information about the maintenance of the behavioral change following the cessation of Tootling.

The most recent Tootling study, conducted by Lum and colleagues (2017), was the first Tootling study to be implemented at the secondary level. The study used an ABAB withdrawal design across three classrooms. Researchers collected class-wide data on both disruptive behavior and academically engaged behavior. Results indicated that the intervention had a moderate to strong effect in reducing disruptive behavior (Tau-U = .78-.94) and a weak to moderate effect in improving academically engaged behavior (Tau-U = .30-.88). This application of Tootling was the first to demonstrate that the intervention can be effective with students at the secondary level; however, no information regarding specific demographics of the students in the class (e.g., grade level, age, socio-economic status) or class size were provided. Furthermore, no treatment acceptability data were collected for participating students. Thus, it is unknown whether this intervention is acceptable to secondary level students.

**Summary of existing Tootling research.** Tootling is a class-wide intervention that combines positive peer reporting with the reinforcement of an interdependent group contingency. The intervention is entirely focused on encouraging and rewarding appropriate student behavior,
and is aligned with the tiered PBIS approach to address student behavioral concerns. Across the six peer-reviewed Tootling studies, results have shown that Tootling increases peer reports of prosocial behavior (Cashwell et al., 2001; Skinner et al., 2000), decreases disruptive behavior (Cihak et al., 2009; Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016), and improves appropriate behavior (Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016). However, some limitations of prior Tootling studies can be noted. First, to date, Tootling has only been implemented in classrooms in the southeastern United States and primarily in rural areas. This limits the applicability of the results to other classroom contexts. Second, only one prior Tootling study involved students at the secondary level and no studies have yet been conducted with middle school students. Additionally, only one study has included any student treatment acceptability assessment; thus it is not clear that children, especially those in secondary school, approve of the intervention. Due to very limited research on class-wide interventions at the secondary level to date, there is little information about which intervention components are effective and acceptable to students and teachers at that grade level. Middle school students are entering puberty, a critical developmental period with significant physical and cognitive changes for students. Thus, determining acceptable and effective interventions for middle school students is even more crucial as students may assert independence from the group intervention or reject interventions as juvenile. Finally, prior Tootling studies have not assessed or programmed for generalization or maintenance of the behavioral impacts from the intervention. Ideally behavioral change resulting from an intervention will endure when an intervention is removed (Freeland & Noell, 2002). Thus, studies are needed that further validate the efficacy and examine the acceptability of the Tootling intervention in different settings and with different student populations.
Conclusion

Significant and prevalent disruptive classroom behavior is interfering with positive academic and behavioral outcomes for a substantial number of students (National Center for Education Statistics, 2012; Public Agenda, 2004). Research has shown that these behaviors are left largely unaddressed (Levitt et al., 2007) and have wide-ranging short and long-term impacts on student mental health and academic achievement (Atkins et al., 2010; Catalano et al., 2004; Levitt et al., 2007). In addition, disruptive student behavior has long been identified by teachers as a primary concern and reason for leaving the filed given its major disruption to instructional time (Emmer & Stough, 2001; Greer-Chase et al., 2002; Public Agenda, 2004; The New Teacher Project, 2013). Although a recent review of teacher preparation programs found that a majority of programs do address classroom management, the focus is not on research-informed practice (Greenberg et al., 2014). Prior to the 1960s, there was limited research-based information on classroom management, and historically both policy and practice have emphasized the use of aversive techniques to manage student behavior (Emmer & Stough, 2001). More recently, classroom management policy and practice have been shifting away from punishment-based strategies to those focusing on promoting and praising positive behaviors (Sugai & Horner, 2006). Research has reflected this shift to positive class-wide behavioral interventions by adapting evidence-based practices that have a more positive focus (Tanol et al., 2010; Wahl, Hawkins, Haydon, Marsicano, & Morrison, 2016; Wright & McCurdy, 2012) and developing new interventions that are aligned with PBIS. Recent meta-analytic work has identified group contingencies as an effective class-wide method at improving student behavior; however overall research is limited, especially in regards to interventions with a focus on appropriate behavior
and those conducted at the secondary level (Chaffee et al., 2017; Maggin et al., 2017). One of the promising new interventions is Tootling (C. H. Skinner et al., 2000), a system of peer-based reporting of positive behavior coupled with an interdependent group contingency. Although there have only been six studies investigating the effects of Tootling, results indicate the intervention reduces classroom disruptive behaviors and improves engagement. Importantly, the intervention is also aligned with PBIS in its focus on acknowledging and reinforcing positive student behavior. Furthermore, Tootling is highly feasible, employing students as agents to recognize positive peer behaviors rather than further taxing stressed teachers. Future research into classroom management interventions should continue to explore Tootling and other class-wide interventions to manage disruptive behavior that are both feasible and positively-focused.
References


https://www.theatlantic.com/education/archive/2015/12/corporal-punishment/420420/


CHAPTER II

Abstract

Class-wide behavioral interventions are a feasible and effective method to support the behavior of all students. In six peer-reviewed studies, Tootling, a class-wide intervention which combines positive peer reporting with an interdependent group contingency, has increased positive peer reports and academically engaged behavior, and decreased disruptive behavior. However, no prior studies have been conducted with middle school students, and none have employed strategies to promote enduring behavior change. An ABAB withdrawal design with maintenance phase, implemented across two middle school classrooms, found moderate effects (NAP = .74; Tau-U = -.48) of Tootling on decreasing disruptive behavior and moderate to large effects (NAP = .76; Tau-U = .68) on academically engaged behavior. Results from the maintenance phase, in which the group contingency was removed, suggest promising strategies to support durable behavioral change. Limitations of the present study, directions for future research, social validity, and implications for practice are discussed.
EFFECTS OF A CLASS-WIDE POSITIVE PEER REPORTING INTERVENTION ON MIDDLE SCHOOL STUDENT BEHAVIOR

Introduction

Teachers consistently rank disruptive student behavior as a primary concern (e.g., Greer-Chase, Rhodes, & Kellam, 2002; Hoglund, Klingle, & Hosan, 2015; Public Agenda, 2004). In recent surveys, over 38% of teachers reported that disruptive behavior interfered with their teaching and over 75% of secondary school teachers believed their teaching would be more effective if these behaviors were reduced (National Center for Education Statistics, 2012; Public Agenda, 2004). Disruptive classroom behavior has been repeatedly related to both immediate and distal negative outcomes for students exhibiting these behaviors including poor academic achievement, delinquency, more significant mental health needs, and adult criminal behavior (Bradley, Doolittle, & Bartolotta, 2008; Greer-Chase et al., 2002; Suldo, Gormley, DuPaul, & Anderson-Butcher, 2014). Exposure to these behaviors has been found to be detrimental to the other students in the classroom, both in terms of lost instructional time and a contagion effect increasing risk of future social-emotional and behavioral issues (Kellam, Ling, Merisca, Brown, & Ialongo, 1998; Sterling-Turner, Robinson, & Wilczynski, 2001). Most critically, the longer disruptive behavior persists, the greater the risk for students and the more intractable, resistant, and expensive it is to treat (Bradley et al., 2008).

In response to the significant school behavioral health need, there have been calls for focusing on prevention, integrating with the public health model, and targeting systems most likely to have the broadest impact (Suldo et al., 2014). One approach aligned with these recommendations involves the use of class-wide behavioral interventions, evidence-based teaching strategies used with all students in a classroom to promote social and behavioral skills
and decrease disruptive behavior (Farmer et al., 2006). In addressing disruptive behavior, class-wide interventions are feasible, provide equity by including all students in the intervention, employ the efficacy of prevention efforts, and have been shown to be effective at improving behavior for all students in the classroom (Chaffee, Briesch, Johnson, & Volpe, 2017).

Unfortunately, however, research has found that the interventions typically used in schools to address disruptive behavior are largely based on personal experience as opposed to empirically validated practices (Bramlett, Murphy, Jonhson, Wallingsford, & Hall, 2002). As such, it is unclear the extent to which the strategies used will actually be effective.

Research on class-wide behavioral interventions has begun to identify specific interventions that general education teachers could use to effectively improve student classroom behavior. Results of a recent meta-analysis of class-wide interventions conducted in the general education classroom indicated that token economies, the Good Behavior Game (GBG; Barrish, Saunders, & Wolf, 1969), and interdependent group contingencies were equally effective at improving problematic student behavior (Chaffee et al., 2017). A majority of the studies identified in the meta-analysis involved group contingencies, in which the consequences delivered (e.g., rewards, punishment) are based on each individual’s behavior (i.e., independent), the collective group performance (i.e., interdependent), or the behavior of a specific student in the group (i.e., dependent; Litow & Pumroy, 1975). Interestingly, however, nearly half of the identified interventions were punitive in nature, such as focusing on problem behavior rather than the appropriate replacement behavior, or involving response cost techniques in which points or rewards are removed rather than earned.

Although there is evidence supporting the efficacy of some of these punitively-oriented interventions (e.g., the GBG; Bowman-Perrott, Burke, Zaini, Zhang, & Vannest, 2016; Chaffee
et al., 2017), punishment-based systems have also been shown to escalate challenging behaviors or incite students to develop covert misbehaviors to avoid punishment (Martin & Pear, 1992; Mayer, 1995; Skinner, 1968). Furthermore, decreases in inappropriate behavior do not always equate to increases in appropriate behaviors. In fact, interventions focusing on teaching or reinforcement of appropriate replacement behaviors have produced more effective and enduring improvements in student behavior than interventions that solely aim to decrease inappropriate behaviors (LeGray, Dufrene, Sterling-Turner, Olmi, & Bellone, 2010; Volmer & Iwata, 1992). Finally, punishment-based interventions stand in marked contrast to recent shifts in research, policy, and practice in school-based discipline towards proactive prevention and increasing appropriate behavior, such as through Positive Behavioral Intervention and Supports (PBIS; Sugai & Horner, 2002).

**Tootling**

One emerging intervention that aims to provide class-wide positive behavior supports is Tootling, a class-wide intervention involving student reports of peer prosocial and appropriate behaviors coupled with an interdependent group contingency reward (C. H. Skinner, Skinner, & Cashwell, 1998). Tootling is a feasible intervention that utilizes peers as monitors and accommodates the realities of teachers who may not be in a position to monitor all student behaviors simultaneously due to many competing stimuli (e.g., focusing on instruction or monitoring a large group; C. H. Skinner, Neddenriep, Robinson, Ervin, & Jones, 2002). In Tootling procedures, the teacher distributes index cards to students and challenges them to record “tootles,” or positive behaviors exhibited by peers (e.g., seeing another student help pick up dropped papers, raising a hand to participate instead of calling out). At the end of the day or period, the teacher collects, counts, and reads aloud five examples of tootles, and then praises the
positive behaviors recorded on the index cards. Through an interdependent group contingency, students in the class all earn a reward when the class as a whole meets its tootling goal. Cumulative progress towards a pre-determined criterion is publicly posted.

Tootling functions by leveraging the praise and positive reinforcement of the teacher, as well as positive peer social pressure, to encourage prosocial behavior from all students. At the core of the Tootling intervention is peer-to-peer written praise, which is aligned with PBIS’ principles of prevention and positive reinforcement of appropriate behavior, and has been identified as a “kernel” or “fundamental unit of behavioral influence” to underlie effective prevention and treatment for children (Embry & Biglan, 2008, p.75). Peer-to-peer written praise has been used successfully in classroom based interventions and school-wide violence prevention to impact student social competence, academic achievement, violence, aggression, physical health, and vandalism (e.g., Embry & Biglan, 2006; Mayer, 1995; Mayer et al., 1993). Social pressure is further leveraged by the interdependent group contingency, as students feel pressure from classmates to help them reach the criterion number of tootles so that they can collectively earn a reward. The visual progress feedback (e.g., thermometer) of cumulative tootles provides additional teacher and social praise for the prosocial behavior (Gresham & Gresham, 1982). Thus, with this class-wide intervention in place, students are likely to exhibit increased rates of positive behavior to increase their individual chances of obtaining peer recognition, teacher praise, and the group reward. With the reciprocal and bidirectional social nature of student behavior in the classroom, the increased rates of appropriate classroom behavior are believed to be naturally reinforcing as they allow for increased positive peer and adult interactions, ultimately facilitating generalization of appropriate behavior (McConnell, 1987). In this manner, the Tootling intervention employs the behavioral process of entrapment,
in which newly acquired social responses come under natural reinforcement (e.g., peer social reinforcement) (Baer & Wolf, 1970).

Over the past 20 years, there have been six published studies examining the effectiveness of Tootling. These studies have shown Tootling produces increased peer reports of prosocial behavior (Cashwell, Skinner, & Smith, 2001; C. H. Skinner et al., 2000), decreased class-wide disruptive behavior (Cihak, Kirk, & Boon, 2009; Lambert, Tingstrom, Sterling, Dufrene, & Lynne, 2015; Lum, Tingstrom, Dufrene, Radley, & Lynne, 2017; McHugh, Tingstrom, Radley, Barry, & Walker, 2016), and increased class-wide engaged behavior (Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016). However, limitations of this work should be noted.

The primary limitation affecting the generalizability of the current body of literature on Tootling is the age of the targeted students. Five of the six published Tootling studies were conducted in elementary classrooms (Cashwell et al., 2001; Cihak et al., 2009; Lambert et al., 2015; McHugh et al., 2016; C. H. Skinner et al., 2000), and only one study to date has been conducted with students at the secondary level (Lum et al., 2017). Tootling has not yet been implemented at the middle school level. Although Tootling has the potential to be an effective strategy at the middle school level, effectiveness is possibly limited by the academic context and developmental characteristics of middle school students. The transition from elementary to middle school is often students’ first experience with rotating classrooms rather than a primary classroom teacher. Students are suddenly required to develop multiple teacher-student relationships as well as navigate varied academic and behavioral expectations across classrooms. Early adolescence is also a unique developmental period (e.g., external and internal pubertal changes, increase in autonomy) in which students are cognitively able to engage in higher-level thinking; however, are also limited by lagging decision-making and executive functioning.
development and declines in motivation for school (Akos, 2002, 2005; Wigfield, Lutz, & Laurel Wagner, 2005). Furthermore, students are actively engaged in identity formation through negotiation of personal interests, school achievement, and social relationships (Wigfield et al., 2005). Within this context, it is unknown whether the Tootling intervention will positively capitalize on this pivotal time of social pressure and self-growth or be rejected by students as juvenile.

There are also additional areas of concern with prior Tootling studies. All prior studies of Tootling have occurred in the same geographic area (i.e., the southeastern United States), and primarily in rural areas. This limits the generalizability of prior results, as it is unclear whether Tootling would be effective in urban or other geographic areas of the United States. Additionally, only one prior Tootling study (McHugh et al., 2016) included a student acceptability measure and only for targeted elementary students. Student acceptability ratings are critical to the success of Tootling, given that the intervention is largely student-initiated and maintained. Furthermore, as research has noted the limited usage of evidence-based interventions, it is important to assess whether the Tootling intervention is acceptable to both teachers and students (Bramlett et al., 2002). Acceptability is particularly important within the middle school context, as students are entering puberty, developing a sense of personal independence, and experiencing uneven cognitive gains, all of which may interfere with the acceptability of Tootling.

Finally, an important goal of a successful behavioral intervention is to achieve enduring behavioral change (Baer, Wolf, & Risley, 1968). Although several prior Tootling studies have recommended further research into the maintenance of the intervention’s positive behavioral effects (Cihak et al., 2009; Lum et al., 2017), it has not been examined to date. Prior Tootling
studies that included a follow-up phase (Lambert et al., 2015; Lum et al., 2017) allowed the teachers to choose whether or not the intervention was continued during that phase. However, no Tootling study to date has employed a strategy to support maintenance of the effects of the intervention from programmed reinforcers (e.g., group contingency reward) to natural reinforcers (e.g., improved peer relationships, teacher praise).

**Purpose of Study**

This study seeks to contribute to the literature by extending and refining the evidence-base for Tootling. In order to meet the Single Case Design Panel guidelines for classification as an evidence-based intervention (i.e., five studies conducted by three independent research groups involving 20 participants; Kratochwill et al., 2010), additional replications of the effects of Tootling are needed. In particular, evidence of the intervention’s effectiveness at the middle school level is needed if it is to be used considered for use with this age group. The purpose of this study was to evaluate the effectiveness and acceptability of Tootling with middle school students in a general-education environment to decrease class-wide disruptive behavior and increase academic engagement. Therefore, the current study sought to answer the following research questions:

1. Does implementation of Tootling lead to improvements in class-wide behavior (e.g., disruptive and academically engaged behavior) in middle school classrooms? This study was the first Tootling study conducted with middle school students and only the second Tootling study with secondary students. We hypothesized that, similar to the prior Tootling study conducted with older students (Lum et al., 2017), we would find moderate effects on academic engagement and disruptive behavior.
2. Do middle school teachers and students find the Tootling intervention acceptable and usable? The only Tootling study conducted with secondary students to date (Lum et al., 2017) did not include any measure of student acceptability. Although the intervention has been found effective and acceptable with younger students (e.g., McHugh et al., 2016), we hypothesized that it might not be as popular with middle school students due to the unique developmental period; however, we hypothesized that teachers would find the intervention acceptable given that it is largely peer-mediated.

3. Does programming common stimuli support the maintenance of positive behavior change? This study examined maintenance effects of the Tootling intervention by programming common stimuli, one of the generalization techniques outlined by Stokes and Baer (1977), in which the same prompting stimuli are maintained while reinforcement is removed. Although enduring behavior change (i.e., generalization across time) of student behavior is a critical aspect of intervention for educators (Baer, Wolf, & Risley, 1968), generalization and sustainable intervention has not been a primary focus of educational research (Freeland & Noell, 2002). Only two prior Tootling studies have included a follow-up phase. Lambert et al. (2015) found that both teachers continued to use the intervention in follow-up and behavioral improvements maintained. However, Lum et al. (2017) found that teachers did not choose to continue using the intervention and the positive behavioral effects began to reverse. No prior Tootling studies have examined any strategies to promote maintenance of the intervention effects. We hypothesized that that during the maintenance phase without the group reward, the positive behavioral effects of the intervention would reverse slightly, but not return to baseline levels.
Method

Permission to conduct the study was obtained from school district administrators and school principals. Additionally, all procedures were approved by Northeastern University’s Institutional Review Board. Informed consent was obtained from each participating teacher.

Participants and Setting

Two general education classrooms from a middle school (grades 6-8) in a northeastern metro area participated in this study. The school had approximately 614 students enrolled with 9.6% receiving free or reduced lunch, 6.5% English Language Learners, and 19.1% students receiving special education support. The middle school used a rotating schedule and had six 42-49 minute blocks within one school day. The two participating teachers had expressed a desire for assistance and were referred by the school psychologist. Classrooms were considered appropriate for the intervention if students in the class were exhibiting disruptive behavior in at least 30% of observed intervals during a 20-minute screening observation (see Procedure).

Classroom A was a sixth grade, general education English/Language Arts classroom consisting of 17 students (7 females, 10 males), one of whom was on a 504 plan (see Table 1). The teacher was a 31-year-old Caucasian female with a Master’s degree in her sixth year of teaching. Classroom B was a sixth-grade general education inclusion Social Studies classroom of 24 students (11 females, 13 males), four of whom received special education services. The teacher of Classroom B was a 54-year-old Caucasian male with a Master’s degree in his thirtieth year of teaching. Classroom B was a special education inclusion class where a special education assistant teacher was present in the classroom three out of every six days. The classes were taught during the same rotating block and observation order alternated when they occurred within the period (e.g., first 20 minutes, last 20 minutes).
Materials

Students were provided with neon colored index cards on which to write their tootles, which were placed at the center of a cluster of 4-5 desks. A plastic container was placed on the teacher’s desk to collect the tootles. Teachers were provided with a goal thermometer to display the students’ collective progress towards the predetermined class goal. This thermometer was clearly visible on the whiteboard of each classroom. A list of possible rewards was developed through consultation with each teacher. Students were also solicited for ideas of rewards and ultimately voted from the list to determine rewards with high preference value. Rewards chosen included students picking their own seating and extra recess.

Dependent Variables

Class-wide behavior. The two primary dependent variables in this study were class-wide disruptive behavior (DB) and academically engaged behavior (AEB). Disruptive behavior was selected as a primary target for intervention because of its association with both immediate and long-term impacts on behavioral and mental health, as well as educational achievement for all students in the classroom (e.g., Greer-Chase et al., 2002; Kellam et al., 1998). Phase change decisions were made based on visual analysis of DB data. As an important academic enabler contributing to academic and classroom success (DiPerna, Volpe, & Elliott, 2002), AEB was selected as a secondary target for intervention.

Per prior Tootling studies (Lambert et al., 2015; Lum et al., 2017), DB was operationally defined as a student demonstrating: (a) out of seat behavior without permission (defined as buttocks not in contact with the seat); (b) audible vocalizations that are not permitted, including talking, singing, whistling, calling out; or (c) motor activity not associated with the assigned task, including physically touching another student or manipulating objects (e.g., playing with paper).
DB did not include quietly raising a hand to answer a question or talking with peers regarding assignments with permission. AEB included both passive and active academic engagement. Active engagement was operationally defined as when the student was actively involved with academic tasks (e.g., reading aloud, writing) and/or speaking with a teacher or peer about the assigned material. Passive engagement was defined as attending to (e.g., looking at, listening to) the assigned work (e.g., independent seatwork, teacher instructions, class-wide activities, group work). AEB did not include calling out or aimlessly looking around the classroom.

Data were collected in each classroom by the primary investigator and two trained observers. Observers were graduate students in school psychology who completed an observation training prior to the implementation of the study. During the training session, observers were trained on the operational definitions for DB and AEB by the primary investigator. Once the operational definitions were mastered, as demonstrated by above 90% accuracy on a quiz (see Appendix B), observers conducted simultaneous observations of previously coded videos until an .80 interobserver agreement criterion was achieved. Interobserver agreement was calculated using the interval-by-interval method (Cooper, Heron, & Heward, 2007), in which the total number of agreements was divided by the total number of intervals, and then multiplied by 100.

The dependent variables were measured during 20-minute observations using a 15-second momentary time sampling and partial interval recording procedure. First, AEB was measured using momentary time sampling, which has been found to be more representative of actual behavior and less error-prone as compared to partial and whole interval recording methods (Radley, O'Handley, & LaBrot, 2015). Using the Behavioral Observation of Students in Schools (BOSS; Shapiro, 2013) mobile application, observers were prompted to observe and record
target behaviors at the end of each interval with a vibration. When prompted, the target student was momentarily observed and the observer recorded whether the student was exhibiting AEB or not. Second, DB was measured using a partial interval procedure. Given that behavior was assessed on a class-wide basis with students displaying varying levels of DB, partial interval recording was more likely to capture instances of DB and prevent floor effects. Following the vibration, the target student was observed until the next cue to determine whether DB occurred or not during that interval. The observer also took notes about the time of the observation, number of students in attendance, and class activities (see Appendix C).

For both recording procedures, the observer rotated to a new student each new interval in a predetermined fixed pattern based on the student seating chart. This pattern was continuously repeated for all intervals in the 20-minute observation. Based on prior research on observational methods, the method of observing engagement in individual students in a fixed pattern provides valid estimates of class-wide behavior and is feasible (Briesch, Hemphill, Volpe, & Daniels, 2015). Data collection procedures were consistent across screening, baseline, intervention, and maintenance phases. Both AEB and DB were reported as percentage of intervals of occurrence, calculated by dividing the total number of intervals of occurrence by the total number of intervals in the observation, and multiplying by 100. Observations were scheduled for all of the school days of the study; however, observations were either not conducted or data were not included if the class’ activity deviated significantly from typical instruction (e.g., state-wide standardized testing, guest speaker, exam).

**Social validity.** In order to more fully understand the acceptability of and satisfaction with the intervention procedures, both teacher and student social validity were assessed using the Usage Rating Profile-Intervention, Revised (URP-IR; Chafouleas, Briesch, Neugebauer, &
Riley-Tillman, 2011) and the Children’s Usage Rating Profile (CURP; Briesch & Chafouleas, 2009a) following the completion of the study. Based on the Health Belief Model (HBM), the URP-IR and the CURP expand beyond treatment acceptability to include internal (i.e., perceived severity of problem, perceived benefits of adherence) and external factors (i.e., anticipated costs, perceived ability to implement the intervention) influencing the usage of a given intervention (Briesch & Chafouleas, 2009b; Briesch, Chafouleas, Neugebauer, & Riley-Tillman, 2013).

**Teacher-reported usability.** Teachers completed a modified URP-IR upon completion of the intervention. The URP-IR is a 29-item scale which assesses multiple factors related to sustained intervention usage including acceptability, teacher understanding, home-school collaboration, feasibility, system climate, and system support (Chafouleas et al., 2011). Items are rated on a 6-point Likert scale with higher ratings generally indicating higher perceived usability. The URP-IR has demonstrated solid internal consistency, with the reported Cronbach’s alpha for the subscales ranging from .67 to .95 (Briesch et al., 2013). Modifications included changing the tense of some words and the addition of Tootling-specific language (see Appendix D). Prior research has found that such minor modifications of wording do not impact the overall psychometric properties of the scale (Freer & Watson, 1999).

**Student-reported usability.** Student usability was assessed using the CURP. The CURP is a 21-item questionnaire on which students are asked to rate the personal desirability, feasibility, and understanding of the intervention using a 4-point Likert scale with higher ratings indicating higher usability. Briesch and Chafouleas (2009b) reported Cronbach’s alpha for the three subscales of the CURP to range from .75 to .92. Similar to modifications to the URP-IR, the tense of some words was changed and the language made more specific to the Tootling
intervention on the CURP (see Appendix E). Adjustments were made to the CURP to ensure that the name of the intervention matched the name chosen by the students in the class.

**Study Design and Procedures**

A single-subject, ABAB reversal design with a maintenance phase was implemented in two middle school classrooms. Reversal designs are the most powerful within-subject design for demonstrating functional relations between the independent and dependent variables through prediction, verification and replication (Gast & Baekey, 2014). Per What Works Clearinghouse (WWC) SCD study design standards, the reversal design includes three demonstrations of the effect of the intervention (i.e., baseline changes as a result of the intervention, behavior reverts to baseline levels when intervention is removed, behavior again changes as a result of the re-introduction of the intervention) at three different points in time (Kratochwill et al., 2010). At least five observations were collected within each phase (Kratochwill et al., 2010). Phase change decisions were made based on visual analysis of DB data, given that this was the primary dependent variable.

**Pre-baseline.** Prior to baseline, teachers were asked to use typical classroom management procedures. Similar to procedures used in prior Tootling studies (Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016), researchers conducted a 20-minute screening observation to ensure that the classroom met the inclusion criteria for students exhibiting class-wide DB in at least 30% of observed intervals (Classroom A = 40.00%, Classroom B = 31.25%). This screening observation was intended to prevent floor effects and ensure that the baseline level of DB was in need of change (Kratochwill et al., 2010).

**Baseline.** Both classes began in the baseline phase, in which data were collected on the dependent variables while typical instructional and classroom management practices were in
place. Teacher A reported using flexible seating, logical consequences, as well as “space and time” as classroom management procedures. Teacher B described his classroom management procedures as “non-existent.” He reported using verbal prompts and writing the names of misbehaving students on the board. Classes remained in the baseline condition for at least five days or until a predictable pattern of behavior was established.

**Teacher training.** Following establishment of a stable pattern of behavior in the baseline phase, the researcher conducted a training session with each teacher. Teachers were provided a script to train students in the intervention, as well as a script to use in the daily implementation of Tootling (see Appendix F). During the meeting, the teachers were also assisted in developing a list of possible class rewards.

**Intervention.** Based on the prior implementation of Tootling at the high school level (i.e., Lum et al., 2017) and other group contingency interventions implemented at the secondary level (e.g., Kleinman & Saigh, 2011; Mitchell, Tingstrom, Dufrene, Ford, & Sterling, 2015), adaptations made to the intervention included (a) describing the intervention as a competition, (b) soliciting names for the intervention from students and then conducting an anonymous vote, and (c) calling the positive peer reports “positive comments” rather than tootles. On the first school day of the intervention, the teachers introduced the Tootling intervention to the class and conducted a student training session. The teacher defined “positive comments” for students and provided both positive and negative examples. Students were also given an opportunity to write sample positive comments and teacher feedback was provided regarding the acceptability of the comments (i.e., name an individual student; specific, anonymous, focused on observed behavior). Each teacher solicited student ideas for names of the intervention and then conducted an anonymous vote. Classroom A chose the name *Mores Magni Challenge*, which is a Latin
translation of Good Behavior Challenge, and Classroom B chose the name Complementation. In addition to the teacher-identified reward list generated during the teacher training, the students in each class provided additional reward suggestions. In a class-wide vote, Classroom A selected a 15-minute recess and Classroom B selected students choosing their own seats for a day.

The school day following the student training, each classroom teacher dispensed 3x5 inch note cards to the students and encouraged students to document their peers’ appropriate behavior (i.e., tootles). Prior to the end of each class, the teacher circulated the classroom with a provided container to collect the written positive comments. Five minutes prior to class dismissal, the teacher randomly selected five tootles, read them aloud to the class, and provided praise to the students for the appropriate behavior listed on the tootle. The teacher also announced how many tootles were reported and recorded the number of tootles on a provided thermometer visual display that documented the class’ progress towards the cumulative goal. Tootles which were incorrect or inappropriate (e.g., identifying inappropriate behavior, a joke, identifying a positive attribute rather than describing a prosocial peer behavior) were ignored and not counted. Multiple tootles reporting the same pro-social behavior were counted individually.

The criterion for earning the interdependent class reward was determined by the teacher based on the value of the selected reward, and was initially set at 50 by each teacher. The teacher announced this criterion to the students in the class. When the class met the goal linked to the chosen reward (e.g., 50 tootles for a 15-minute recess), the entire class received the predetermined reward. If the class did not meet the goal, the total number of tootles was applied to the next day’s total. Following the second day of the intervention, however, anecdotal reports suggested that students in Classroom B began openly discussing how to “outsmart” the intervention and earn the reward immediately by writing three positive comments each, often of
duplicate behaviors. Upon consultation with the experimenter, Teacher B adjusted the procedures to distribute only two index cards to each student and increased the criterion for reward to 160 positive comments.

**Withdrawal.** Following clear treatment effects on DB in each classroom during the first intervention phase, the positive peer reporting procedures and interdependent group contingency were withdrawn. The teachers notified students that the class would no longer be playing the game and all intervention materials were removed. As in the baseline condition, observers conducted daily observations of class-wide DB and AEB.

**Reimplementation.** After the withdrawal phase, the positive peer reporting procedures and group contingency were reimplemented as they were in the initial intervention phase. Both teachers chose to reinstate the classroom’s progress towards their previous goal.

**Maintenance.** Once a treatment effect was documented in the second intervention phase, teachers were instructed to move into the maintenance phase. In Tootling procedures, positive behavior is encouraged and reinforced by peer recognition of appropriate behavior, teacher verbal praise, and the interdependent group contingency. In order to promote the maintenance of the behavioral effects of the Tootling intervention, a technique called programming common stimuli was used. Programming common stimuli is the process of providing students with “sufficient stimulus components occurring in common in both the training and generalization settings” (Stokes & Baer, 1977, p. 360). Previous studies, especially within the field of applied behavior analysis, have ensured that generalization occurred by utilizing multiple physical stimuli (e.g., desks, chairs), people (e.g., teachers, peers), or training procedures (e.g., salient visual stimuli; Mesmer, Duhon, & Dodson, 2007; Stokes & Baer, 1976, 1977) during training procedures. In this study, the procedures of peers recording the tootles and teachers reading them
aloud remained constant, but the programmed consequences of the interdependent group contingency were removed. This condition lasted five days. Teacher and student social validity assessments (e.g., URP, CURP) were administered following the last day of the maintenance phase.

**Interobserver agreement.** Per single-case design standards from the WWC (Kratochwill et al., 2010), interobserver agreement (IOA) for class-wide DB and AEB was assessed between the primary research assistant and another trained observer for a minimum of 33% of observations (Classroom A: 46% of all observations, Classroom B: 38% of all observations) for each phase of the study in each classroom. IOA data were calculated as described earlier. Observers were required to obtain at least 80% IOA with the primary observer. If agreement fell below that criterion, observers were retrained on the procedures and operational definitions before conducting further observations. This retraining occurred only once for Classroom A (day 3 of the second intervention phase). IOA for Classrooms A and B averaged 91.97% (Range = 78.75-100%) and 92.50% (Range = 85.00-98.75%), respectively, across both DB and AEB.

**Treatment integrity.** Treatment integrity was assessed throughout the intervention, withdrawal, and maintenance phases in order to ensure and verify that the teachers and students were trained appropriately and the intervention was implemented as planned. Treatment integrity measures included checklists to monitor the teacher training conducted by the primary investigator and student introduction and intervention training by the teacher (see Appendices G, H). Integrity data reflected 100% fidelity to procedures across the two teachers during both the teacher Tootling training sessions and the student Tootling training sessions.

The primary observer collected treatment integrity data for 100% of observations during intervention, withdrawal, and maintenance phases (see Appendices I, J, K). Treatment fidelity
IOA was obtained by a secondary observer for at least 33% (Classroom A: 48%, Classroom B: 38%) of occasions. Observers measured intervention treatment integrity in both classrooms using a 10-item checklist. Treatment integrity, as rated by the observers, averaged 96.15% (range = 80.00–100%) of steps completed for Classroom A and 93.33% (range = 70.00-100%) for Classroom B. Although levels of treatment integrity were high overall, it is of note that at times both teachers failed to implement the same items, which included reminding students to be observing for positive peer behaviors, reviewing the procedures, and reminding students of their progress towards the goal (see Table 2). The teachers may have felt that these components of the procedure were redundant on a daily basis. Treatment integrity data during withdrawal and maintenance phases averaged 100% in both classrooms. IOA for treatment integrity was 100% across all observations in both classrooms.

Additionally, teachers were provided with self-monitoring checklists of daily intervention and maintenance procedures (see Appendices L, M, N); however, neither teacher regularly completed these checklists. Instead, both teachers reported primarily using the checklists as a prompting tool.

Research has shown that performance feedback is a critical component of improving implementation of an intervention (Codding, Livanis, Pace, & Vaca, 2008). Performance feedback was given to teachers by the primary investigators if treatment integrity data fell below 80%, or if the teacher sought assistance. The only instance of performance feedback occurred for Teacher B on the second day of the first intervention phase, when treatment integrity was 70% and the teacher sought assistance with an insincere student response to the intervention.
Data Analysis

Class-wide behavioral data were graphed as a percentage of intervals in which DB and AEB were recorded out of total intervals during 20-minute observations. The percentage of intervals in which the students displayed DB and AEB was calculated separately by dividing the total number of intervals of occurrence by the total number of intervals in the observation and multiplying by 100. Although visual analysis may examine several different features of the data (e.g., changes in level, trend, and variability, as well as the immediacy of effect), data gathered for both classrooms during baseline and intervention were inspected visually for changes in level and trend to determine treatment effects (Kazdin, 1982). Trend and level were emphasized during visual analysis as opposed to immediacy of effect because Lum et al. (2017) showed a gradual overall behavioral response when implementing Tootling with older students. Although there is natural variability in student behavior due to classroom factors (e.g., academic demand, instructional modality), both a gradual directional and average overall change is expected in response to the introduction and withdrawal of the intervention. Given the lack of consensus regarding effect size analytics for single-case design studies the WWC recommends the use of multiple effect size indices (Kratochwill et al., 2010). Thus, two effect size calculations, Nonoverlap of All Pairs (NAP; Parker & Vannest, 2009) and Tau-U (Parker, Vannest, Davis, & Sauber, 2011) were used to further evaluate the treatment effects of the Tootling intervention. NAP and Tau-U have been used as effect size indices in prior Tootling studies (e.g., Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016), allowing for direct comparison of the effects of Tootling on student behavior. All NAP and Tau-U calculations were performed for each adjacent phase contrast separately (i.e., each A-B of the withdrawal design) and then combined into an overall weighted effect size for each classroom using a web-based calculator (Vannest,
Parker, Gonen, & Adiguzel, 2016). Data from the maintenance phase were not included in the overall weighted effect size. Maintenance data were compared to the initial baseline phase as an index of the durability of behavioral change (Beeson & Robey, 2006).

**NAP.** The NAP effect size metric builds upon the visual analysis of single-case design graphs and, like other non-overlap metrics (e.g., Percentage of Non-overlapping Data (PND), Percentage of All Non-Overlapping Data (PAND), Percentage Exceeding the Mean (PEM)), summarizes data overlap between phases A and B. Related to Area Under the Curve (AUC) analyses, NAP compares each phase A data point with each phase B data point and determines the probability that a random treatment data point will be greater than a random baseline data point. Although NAP is impacted by baseline trend, it has no a priori assumptions, includes all data points in calculations, and has shown good discriminability as compared to other nonoverlap metrics (Parker & Vannest, 2009). NAP values between .00 and .65 are considered weak effects, scores between .66 and .92 are moderate effects, and scores from .93 to 1.00 are considered strong or large effects (Parker & Vannest, 2009).

**Tau-U.** Tau-U is a conservative nonoverlap effect size metric that allows for baseline and/or intervention phase trend control. As it includes all data points in calculations, it is resistant to outliers and a small number of data points. Tau-U is derived from the Kendall’s Rank Correlation and the Mann-Whitney U test between groups. In contrast to methods relying on parametric assumption or linear trends, Tau-U is more reliable at identifying trend with limited data points (Parker et al., 2011) and has been used recently within the single-case literature (e.g., Bowman-Perrott et al., 2016; Chaffee et al., 2017) and in prior Tootling intervention studies (Lum et al., 2017). However, like all single-case effect sizes, Tau-U has limitations. Tau-U is affected by the number of data points, values may exceed the conventional bounds of +/-1, and it
is not sensitive to the magnitude of change when there is no overlap between baseline and intervention (Tarlow, 2017). Despite these limitations, Tau-U was selected for the aforementioned advantages and to allow for direction comparisons of the results of this study with prior Tootling and class-wide intervention studies. Tau-U values of .20 have been interpreted as a small effect, .20-.60 as a moderate effect, .60-.80 a large effect, and above .80 as a large to very large effect (Vannest & Ninci, 2015).

Results

The Tootling intervention was evaluated in two middle school classrooms using a single-case reversal design. Experimental control was demonstrated in Classroom A, with increases in academic engagement and decreases in disruptive behavior observed each time that the intervention was introduced. However, the results obtained in Classroom B are more difficult to interpret, as some threats to internal validity occurred. Results specific to each classroom are discussed in the following section.

Classroom A

During baseline, the students in Classroom A exhibited DB during an average of 26% of intervals (range = 19-31%), with no visible trend in the data (Figure 1, top panel). When Tootling was introduced, the mean level of DB decreased to 17% (range = 4-24%), and a significant decreasing trend was observed. The intervention was then withdrawn and the level of DB immediately increased from 4 to 15%. Throughout the withdrawal phase, the level of DB remained fairly consistent (Mean = 15%; range = 13-21%) and lacked trend. When the intervention was re-introduced, the level of DB decreased slightly (Mean = 12%; range = 7-16%) with an overall slight decreasing trend. During the maintenance phase, DB was observed during an average of 15% of intervals (range = 10-24%) and no trend was observed in this phase.
Across the intervention phases, the data show a strong or slight negative trend. Overall, Tootling in Classroom A had a moderate effect (Tau-U = -.48; NAP = .92) in decreasing DB when using weighted Tau-U and NAP calculations (see Table 3). As an index of durable change from the intervention, DB in baseline compared to the maintenance phase demonstrated a moderate (NAP = .92) to large effect (Tau-U = -.84)

During the baseline phase, class-wide AEB data reflect a negative trend and occurred in an average of 74% of intervals (range = 69-83%). However, when Tootling was introduced, AEB displayed a strong, positive trend coupled with increases in level (Mean = 84%; range = 75-99%). When Tootling was subsequently removed, the class-wide AEB level immediately dropped from 99 to 71% and showed a decreasing trend (Mean = 80%; range = 70-86%). The intervention was then re-introduced and the level increased (Mean = 84%; range = 75-95%), and an increasing trend was observed. Finally, during the maintenance phase AEB decreased slightly in level (Mean = 80%; range = 76-86%) with no visible trend. Both intervention phases exhibited strong positive trends and levels of AEB that approached the ceiling (100%), potentially limiting a full assessment of the effects of the intervention. Overall, calculations indicated Tootling had a moderate (NAP = .76) to large (Tau-U = .68) positive effect on AEB. Analyses of maintenance phase data compared to baseline showed a moderate (NAP = .80) to large (Tau-U = .76) durable effect of Tootling on AEB.

**Classroom B**

During the baseline phase, students in Classroom B (Figure 1, bottom panel) demonstrated class-wide DB during an average of 37% of intervals (range = 23-45%) and a slight increasing trend was observed. When Tootling was introduced, the level of DB decreased substantially to an average of 20% of intervals (range = 20-34%) with no visible trend observed.
Upon withdrawal of the intervention, DB initially increased but then decreased on the third day of implementation (at which time a staffing change occurred), leading to a steep negative trend (Mean = 20%; range = 13-31%). The intervention was reintroduced, and the level of DB initially increased from 12 to 25%; however, the overall level was equivalent to that observed during the withdrawal phase (Mean = 20%; range = 14-28%). The data in this phase show a slight decreasing trend. Finally, during the maintenance phase DB decreased in level (Mean = 14%; range = 5%-24%) with a decreasing trend. Although effect size calculations for class-wide DB in Classroom B indicated an overall weak effect (NAP = .63; Tau-U = -.21), results suggested a large durable effect when comparing the baseline to maintenance phase (NAP = .96; Tau-U=-1.04; see Table 3).

Class-wide AEB averaged 65% of intervals (range = 56-75%) with no visible trend observed during baseline. When the intervention was introduced, the level of AEB increased (Mean = 75%; range = 68-80%), with no trend observed. Upon the withdrawal of Tootling, the intervals in which AEB occurred initially decreased from 78 to 65%; however, on the third day of this phase, the level of AEB increased dramatically from 56 to 79%. This dramatic change in behavior coincided with the above-mentioned staffing change. Overall, AEB averaged 72% of intervals (range = 56-79%) during the withdrawal phase with an overall increasing trend observed. Tootling was then reintroduced and there was an increase in the level of AEB (Mean = 82%; range = 74-89%) with a slight increasing trend seen. During the maintenance phase AEB remained at the same level as the prior intervention phase (Mean = 82%; range = 73-89%) with a slight increasing trend visible. Tau-U calculations indicated Tootling had a moderate (NAP = .92) to large (Tau-U = .79) effect at increasing AEB in Classroom B. The durable effects
(Baseline v. Maintenance) of Tootling on AEB for Classroom B were large (NAP = .96; Tau-U = .92) for both effect size calculations.

Social Validity

*Teacher-reported usability.* Results of the URP-IR (see Table 4) suggest that both Teacher A and Teacher B understood the intervention components and believed that the intervention was feasible and that it required minimal home-school collaboration. However, Teacher A reported higher levels than Teacher B on the factor of Acceptability. Specifically, Teacher B slightly disagreed and disagreed that the intervention was a *fair way* or a *good way* to handle the child’s behavior problem. Teacher B’s results also indicated that he believed that the intervention required significant system supports (e.g., consultative support, professional development); however, Teacher A’s ratings indicated that the intervention could be implemented with low levels of system supports.

*Student-reported usability.* CURP results (see Table 5) indicated that whereas both classes agreed that the intervention was feasible and that they understood the intervention, students in Classroom A reported somewhat higher levels of Personal Desirability (Classroom A Mean = 3.20; Classroom B Mean = 2.80). For example, in response to the statement *I like this intervention,* 82.4% of students in Classroom A agreed, whereas only 77.3% of students in Classroom B agreed.

Discussion

Although there is an emerging evidence base in support of the use of Tootling to decrease disruptive classroom behavior (e.g., Cihak et al., 2009; Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016) and increase appropriate behavior or academic engagement (e.g., Lambert et al., 2015; Lum et al., 2017; McHugh et al., 2016), studies to date have implemented these
procedures almost exclusively with elementary school students. Use of a positive peer reporting intervention in middle school classrooms has the potential to capitalize on this pivotal time of social pressure and self-growth; however, at the same time the intervention may be rejected as students attempt to assert independence from adults as well as their own maturity. This study sought to replicate the effects of prior studies of Tootling through an ABAB single case reversal design in two middle school classrooms. Results show via visual analysis that experimental control was documented in Classroom A. In Classroom B, however, potential threats to internal validity in the form of a staffing change, may have prevented the establishment of experimental control and three clear replications of effect. Overall, based on both visual and quantitative analyses, there is one convincing and one more cautious demonstration of an effect of Tootling with middle school students.

Additionally, this study sought to extend the research by assessing both teacher and student ratings of the usability of the intervention. As hypothesized, results indicated that middle school students found the intervention slightly less acceptable than elementary students; however, overall levels of acceptability were still high. Finally, this study sought to explore the effectiveness of programmed common stimuli to facilitate durable behavioral change. We found that the generalization strategy is a promising method to extend the positive effects of the intervention. These results are discussed in greater detail below.

**Effect of Tootling on Class-wide Behavior**

Visual analyses of both DB and AEB in Classroom A demonstrated experimental control via three replications of the effect of Tootling. However, although there was a response in the expected direction, neither DB nor AEB fully returned to baseline levels during the withdrawal phase. Although this carryover is not desired with regards to experimental control, it may
indicate that Tootling helps students learn to display positive behavior even without the external reward in place. Regarding DB, quantitative analyses utilizing both NAP and Tau-U found moderate effects of Tootling on DB (NAP = .74; Tau-U = -.48). These results are slightly smaller than the effects found by prior published Tootling studies (i.e., Lambert et al., 2015: NAP = .88-1.00; McHugh et al., 2016: NAP = .92-1.00; Lum et al., 2017: NAP = .92-.95). Additionally, both NAP and Tau-U analyses found moderate to large effects of Tootling on AEB (NAP = .76; Tau-U = .68). These results replicate the effects found by prior elementary-level studies (i.e., Lambert et al., 2015: NAP = .90-1.00; McHugh et al., 2016: NAP = .92-1.00), and are slightly larger than effects previously found at the high school level (i.e., Lum et al., 2017: NAP = .82-.86) for AEB. As hypothesized, students experiencing the unique developmental period of middle school may be slightly less receptive to the components of Tootling, potentially due to peer social pressure around positive comments. Students may consider it the teachers’ role to dispense compliments and thus may resist writing and receiving positive comments themselves for fear of being aligned with the adults rather than their peers. Alternatively, or in conjunction, during this period of self-identification that occurs in the middle school years, students may be self-conscious about appearing too juvenile in appreciating the positive comments. However, given that Tootling is largely student-maintained and a feasible class-wide intervention, the moderate to large effects sizes on class-wide DB and AEB are encouraging for general education teachers seeking Tier 1 interventions for middle school classrooms.

Although experimental control was established in Classroom A, the data in Classroom B are less conclusive. Effect size comparisons of baseline to the first Tootling phase indicate moderate to large effects for DB and large effects for AEB. These results were present despite the teacher’s decision to substantially increase the criterion for reward (i.e., from 50 to 160
positive comments) and limit students to two positive comments each per class session. These changes were made by Teacher B to maintain the acceptability of the intervention when students started to “game” the positive comment system, and served to distance the reward. Although replication of the effects appeared to continue- at least initially- in the withdrawal phase, there was a clear shift in the levels of both AEB and DB from the second to the remaining three data points in the phase. Specifically, DB dramatically decreased in level and stabilized, and AEB correspondingly increased in level, as well. As this shift coincided with the return of an assigned paraprofessional from medical leave and the multi-day absence of a particularly disruptive student, it is unknown whether students’ subsequent behavioral response is due to the intervention, the changes in the classroom environment, or both. It should be noted that the levels of target behaviors seemed to stabilize in the withdrawal phase following the return of paraprofessional, and student behavior in Classroom B responded to the re-implementation of Tootling with increases in AEB and a decreasing trend in DB. However, as neither experimental control nor the WWC standards were established for Classroom B, these promising results are inconclusive and must be interpreted with caution.

**Strategies for Generalization of Behavioral Change**

In order to demonstrate meaningful change in behavior, behavioral interventions must achieve results that endure over time; however, this has not been examined in prior Tootling studies. In fact, only two prior Tootling studies included a follow-up phase (Lambert et al., 2015; Lum et al., 2017). One study found that teachers continued to employ the intervention (Lambert et al., 2015); however, the other study found that after one week teachers were no longer using any intervention components and class-wide behaviors had nearly returned to baseline levels (Lum et al., 2017). To address this limitation present in previous studies, the current study
explored the use of a specific strategy—programming common stimuli—to maintain the behavioral gains of the Tootling intervention. In line with our hypothesis, we found that behavioral changes from Tootling were durable when positive comments procedures were maintained, even with the removal of the interdependent group contingency. This, along with the carryover in the withdrawal phase, suggests that the promotion of a positive classroom environment through Tootling may lead to entrapment, the shift from reliance on external rewards to the natural reinforcement of the teacher and peer social recognition. In fact, Classroom A’s index of durable change (Baseline v. Maintenance phases) for both DB and AEB indicated moderate to large effects. Though the data should be interpreted with caution, visual analyses of Classroom B’s maintenance phase seem to confirm, or even improve upon, the maintenance results of Classroom A with the index of durable change suggesting large effects across both effect size metrics and dependent variables (i.e., DB and AEB). These results indicate that it may not be necessary to keep the rewards component of Tootling in order to continue to see positive behavioral effects. This is particularly important as research has shown teachers have difficulty sustaining intervention components (Han & Weiss, 2005) and may find external reinforcement procedures less acceptable (Akin-Little, Eckert, Lovett, & Little, 2004).

Social Validity

Even with evidence of effectiveness and strategies to support durable behavioral change, an intervention is of limited value if teachers and students do not want to actually use it. Based on prior studies, we hypothesized that teachers would find the intervention acceptable, but that students might have slightly lower ratings of acceptability given the identity development of middle school students. Teachers’ ratings suggest that they both understood the intervention components, believed that the intervention was feasible, and felt that the intervention required
minimal home-school collaboration; however, Teacher B found the intervention less acceptable (e.g., felt it was not fair to manage student behavior using Tootling) as compared to Teacher A. As was previously discussed, according to anecdotal report, Teacher B’s baseline classroom management practices emphasized punitive individual student consequences. Research on teacher intervention acceptability has found higher levels of acceptability if the intervention is aligned with teachers’ professional and philosophical orientation (Witt, Elliott, & Martens, 1983). Thus, the fact that the collective and positive focus of the Tootling intervention conflicted with Teacher B’s typical classroom management style likely contributed to these lower ratings.

Although teacher acceptability should be considered, as a primarily peer-maintained intervention, it is particularly important that students view the Tootling intervention as acceptable and feasible. Student results for each classroom paralleled the teacher ratings. Students in Classroom B had lower ratings of personal desirability of the intervention, as compared to students in Classroom A, although the majority of students in both classes reported liking the intervention. Teacher treatment acceptability directly relates to the degree and quality of intervention implementation (Wolf, 1978), which in turn could impact the students’ acceptability of the intervention. Although this study reported high levels of integrity across the training of teachers, students, and implementation across phases, the quality or enthusiasm of the intervention implementation was not formally assessed. However, informal observations by research assistants noted Teacher A’s greater enthusiasm for the intervention, as compared to Teacher B. Thus, the lower student ratings in Classroom B may be a reflection of Teacher B’s lower acceptability of Tootling.
Limitations and Directions for Future Research

As is the case with any study conducted in an applied setting, there are some aspects that could not be controlled and may have impacted the results. Of primary consequence in this study, the planned replication of Classroom B was inconclusive due to the return of a paraprofessional from medical leave during the withdrawal phase. This irreversible event threatened the internal validity of the results. In addition, there are some factors that may limit the generalizability of the results. For example, Classroom A happened to have a relatively small class size (n= 17) compared to Classroom B (n = 24), and did not include any special education students. Thus, with only one conclusive effect, further replications of Tootling at the middle school level with diverse populations are needed to enhance the external validity and generalizability of the results.

There were also some limitations with regards to study design that may have prevented the demonstration of strong experimental control or a full examination of the effects of the intervention. First, Classroom A demonstrated ceiling effects with AEB during both intervention phases. Although a screening observation standard of 30% of intervals of DB was administered to prevent floor effects, future studies may consider a stricter standard or one that also involves a criterion for AEB. It may also be beneficial for future studies to employ a multiple baseline design, as this would allow for a longer intervention phase to more fully examine the effects across multiple reward cycles and address any entrapment effects during the withdrawal phase. Additionally, as both classrooms occurred in the same period, observers were not able to directly assess all aspects of treatment integrity on each day of the study as certain components occur at the beginning (e.g., review positive commenting procedures) or end (e.g., read aloud five positive comments) of the class session. Self-report measures were planned to capture these
missing treatment integrity data; however, teachers did not complete them with regularity. The observations of the classrooms that did occur, as well as the teachers’ informal comments, serve as an indication that the intervention was implemented with integrity.

Additionally, although assessed levels of treatment integrity were high, informal assessments of enthusiasm or quality of intervention implementation appeared to differ across classrooms. Dane and Schneider (1998) identified quality of delivery as one of five aspects of treatment integrity that may influence the overall impact of the intervention and validity of the study; however, this study did not include any formal measure of quality of implementation. Future Tootling and class-wide intervention studies should explore the degree to which quality and enthusiasm of implementation impacts the overall effectiveness of Tootling.

Given the positive effects of Tootling that have been noted for academically engaged and disruptive student behavior, future research should continue to evaluate the class-wide intervention. In particular, additional replications are needed at the middle and high school level, an age group with a limited evidence-base. Additionally, based on the durable change sustained in the maintenance phase without the group contingency reward, future studies should compare the effectiveness of Tootling with and without the group contingency component in place.

Finally, the goal of any behavioral intervention is enduring behavioral change. Results from this study suggest the use of programmed common stimuli may promote durable behavioral change; however, the maintenance phase only lasted five days. Future research should evaluate the longer-term effects of Tootling on class-wide behavior change. Additional methods to sustain behavior change should also be explored, such as systemically fading out the interdependent group contingency by gradually increasing the criterion for the reward.
Implications for Practice

In addition to identifying additional avenues for future research, the current study has several implications for applied practice. This and prior studies suggest that Tootling is an effective and feasible class-wide intervention for decreasing overall disruptive behavior and increasing academic engagement. At the core of the Tootling intervention is peer-to-peer written praise, which is aligned with PBIS’ principles of prevention and positive reinforcement of appropriate behavior and allows for the intervention to be largely peer-maintained. Tootling may be particularly appealing to teachers, who are managing the inordinate instructional and behavioral management demands of a classroom, as students are responsible for the majority of tasks (i.e., writing tootles) associated with the intervention. Additionally, as there has been markedly less research conducted on class-wide interventions at the middle and high school level (Chaffee et al., 2017; Maggin, Pustejovsky, & Johnson, 2017), these results will help guide middle school teachers in selecting interventions to support student behavior. Finally, as many evidence-based interventions tend to be focused on negative behaviors (Chaffee et al., 2017), Tootling is a promising option for teachers working within a school-wide PBIS context and seeking an aligned class-wide intervention.
References


*Journal of Behavioral Education, 17*, 4-23. doi: 10.1007/s10864-007-9058-6


*Psychology in the Schools, 39*, 327-335. doi: 10.1002/pits.10022


http://www.publicagenda.org/


Table 1

Demographic Data for Each Participating Classroom

<table>
<thead>
<tr>
<th>Gender</th>
<th>Classroom A</th>
<th>Classroom B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>Classroom A</th>
<th>Classroom B</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>Black</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Asian</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Special Education/504</th>
<th>Classroom A</th>
<th>Classroom B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other Health Impairment</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Traumatic Brain Injury</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>504 Plan</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
### Table 2

*Overall Percentages of Observer-Scored Treatment Integrity*

<table>
<thead>
<tr>
<th>Intervention Step</th>
<th>Classroom</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feedback chart visible</td>
<td>100% 100%</td>
</tr>
<tr>
<td>2. Positive comment box visible</td>
<td>100% 100%</td>
</tr>
<tr>
<td>3. Index cards on students’ desks</td>
<td>100% 100%</td>
</tr>
<tr>
<td>4. Remind students to look out for positive behaviors.</td>
<td>90% 71%</td>
</tr>
<tr>
<td>5. Review positive comments procedures</td>
<td>90% 71%</td>
</tr>
<tr>
<td>6. Discuss cumulative progress towards goal</td>
<td>90% 71%</td>
</tr>
<tr>
<td>7. Read at least 5 positive comments aloud</td>
<td>100% 100%</td>
</tr>
<tr>
<td>8. Praise the behaviors on the positive comments</td>
<td>100% 100%</td>
</tr>
<tr>
<td>9. Record new total positive comments on the feedback chart</td>
<td>100% 100%</td>
</tr>
<tr>
<td>10. Reward the class if they met the goal</td>
<td>100% 100%</td>
</tr>
<tr>
<td></td>
<td>Classroom A</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>NAP</td>
</tr>
<tr>
<td><strong>Disruptive Behavior</strong></td>
<td></td>
</tr>
<tr>
<td>Baseline/Initial Tootling</td>
<td>.78 (M)</td>
</tr>
<tr>
<td>Withdrawal/Reimplementation</td>
<td>.70 (M)</td>
</tr>
<tr>
<td>Weighted Average</td>
<td>.74 (M)</td>
</tr>
<tr>
<td>Baseline/Maintenance</td>
<td>.92 (M)</td>
</tr>
<tr>
<td><strong>Academically Engaged Behavior</strong></td>
<td></td>
</tr>
<tr>
<td>Baseline/Initial Tootling</td>
<td>.84 (M)</td>
</tr>
<tr>
<td>Withdrawal/Reimplementation</td>
<td>.68 (M)</td>
</tr>
<tr>
<td>Weighted Average</td>
<td>.76 (M)</td>
</tr>
<tr>
<td>Baseline/Maintenance</td>
<td>.80 (M)</td>
</tr>
</tbody>
</table>

*Note. NAP = Nonoverlap of All Pairs. * = wrong direction, W = weak effect, M = moderate effect, L = large effect*
Table 4

*URP-IR Mean (SD) by Subscale and Classroom*

<table>
<thead>
<tr>
<th>Classroom</th>
<th>Acceptability</th>
<th>Understanding</th>
<th>Home-school</th>
<th>Feasibility</th>
<th>System climate</th>
<th>System support</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4.78 (0.67)</td>
<td>5.67 (0.58)</td>
<td>1.33 (0.58)</td>
<td>4.83 (0.98)</td>
<td>4.20 (1.30)</td>
<td>1.33 (0.58)</td>
</tr>
<tr>
<td>B</td>
<td>3.78 (0.97)</td>
<td>5.67 (0.58)</td>
<td>1.00 (0.00)</td>
<td>4.33 (1.03)</td>
<td>4.60 (0.55)</td>
<td>4.67 (0.58)</td>
</tr>
</tbody>
</table>
## Table 5

*CURP Mean (SD) by Subscale and Classroom*

<table>
<thead>
<tr>
<th>Classroom</th>
<th>Personal Desirability</th>
<th>Feasibility</th>
<th>Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3.20 (0.31)</td>
<td>1.58 (0.30)</td>
<td>3.35 (0.14)</td>
</tr>
<tr>
<td>B</td>
<td>2.80 (0.26)</td>
<td>1.69 (0.47)</td>
<td>3.34 (0.33)</td>
</tr>
</tbody>
</table>
Figure 1

*Effects of Tootling on Middle School Classrooms*

*Note*: * indicates days on which the class earned the reward.
Appendix A

Tootling Intervention—Demographic Form

**Teacher Information**

1. Age: ________

2. Gender: ______________________

3. Race:   
   - American Indian/Alaskan Native □  
   - Asian □  
   - Native Hawaiian or other Pacific Islander □  
   - Black or African American □  
   - White □  
   - Biracial □

4. Ethnicity:   
   - Hispanic/Latino □  
   - Non-Hispanic/Latino □

5. What courses do you currently teach? ________________________________

6. What grade level(s) do you currently teach? ________________

7. How many years have you been a teacher? _________________

8. How many years have you taught the grade level that you currently teach? ____________

9. Please select your highest degree earned:   
   - High School Diploma □  
   - B.A./B.S. □  
   - M.A./M.S./Ed.M. □  
   - PsyD./Ph.D./Ed.D. □

**Class Information**

1. Class Subject:_____________________________

2. Number of students in the class: _____________

3. Gender:   
   - Male ____  
   - Female ____  
   - Other_____  
   - If Other, please specify: _______

4. Grade: ______________________

5. Race:   
   a. Black or African American_______  
   b. White_______  
   c. Asian_______  
   d. American Indian/Alaskan Native____
e. Native Hawaiian or other Pacific Islander _____

6. Ethnicity: Hispanic/Latino _____ Non-Hispanic/Latino _____

7. How many students are identified as 504 students? ________________

8. How many students are identified as special education students? ________________
   a. Please list the categories under which the student(s) qualify for services? ________________

9. Describe your typical classroom management procedures:
Appendix B

Quiz on Operational Definitions of Dependent Variables

Directions: Based on the definitions below, please identify the following student behaviors as disruptive behavior (DB), academically engaged behavior (AEB), or neither.

<table>
<thead>
<tr>
<th>Disruptive Behavior</th>
<th>A student demonstrating: (a) out of seat behavior without permission (defined as buttocks not in contact with the seat); (b) audible vocalizations that are not permitted including talking, singing, whistling, calling out; and (c) motor activity not associated with the assigned task, including physically touching another student, manipulating objects (e.g., playing with paper). Non-examples: quietly raising a hand to answer a question or talking with peers regarding assignments with permission.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academically Engaged Behavior</td>
<td>AEB includes both active and passive engagement. Active: student is actively involved with academic tasks (e.g., reading aloud, writing) and/or speaking with a teacher or peer about the assigned material). Passive: student is attending to (e.g., looking at, listening to) the assigned work (e.g. independent seatwork, teacher instructions, class-wide activities, group work). Non-examples: calling out or aimlessly looking around the classroom</td>
</tr>
</tbody>
</table>

1. Student tapping pencil on desk. ____________
2. Student quietly completing worksheet independently. ____________
3. Student talking with classmate during independent work. ____________
4. Student staring out into space at the wall. ____________
5. Student doodling on a worksheet paper during independent work. ____________
6. Student singing to him/herself. ____________
7. Student sleeping at desk. ____________
8. Student raises hand and waits quietly to be called on to answer a teacher’s question. ____________
9. Student calls out when teacher asks a question. ____________
10. Student talking with classmate about assignment during group work. ____________
Quiz on Operational Definitions of Dependent Variables Key

*Directions: Based on the definitions below, please identify the following student behaviors as disruptive behavior (DB), academically engaged behavior (AEB), or neither.*

<table>
<thead>
<tr>
<th>Disruptive Behavior</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
</table>
| A student demonstrating: (a) out of seat behavior without permission (defined as buttocks not in contact with the seat); (b) audible vocalizations that are not permitted including talking, singing, whistling, calling out; and (c) motor activity not associated with the assigned task, including physically touching another student, manipulating objects (e.g., playing with paper). Non-examples: quietly raising a hand to answer a question or talking with peers regarding assignments with permission. |                                                                                                                                                                                                           | 1. Student tapping pencil on desk. DB  
2. Student quietly completing worksheet independently. AEB  
3. Student talking with classmate during independent work. DB  
4. Student staring out into space at the wall. Neither  
5. Student doodling on a worksheet paper during independent work. Neither  
6. Student singing to him/herself. DB  
7. Student sleeping at desk. Neither  
8. Student raises hand and waits quietly to be called on to answer a teacher’s question. Neither  
9. Student calls out when teacher asks a question. DB  
10. Student talking with classmate about assignment during group work. AEB |
Appendix C

Tootling Study Observation Form

**Disruptive Behavior:** a student demonstrating: (a) out of seat without permission (defined as buttocks not in contact with the seat); (b) audible vocalizations that are not permitted including talking, singing, whistling, calling out; or (c) motor activity not associated with the assigned task, including physically touching another student, manipulating objects (e.g., playing with paper).

**Academically Engaged (AE) Behavior:** AEB includes both active and passive engagement. **Active:** student is actively involved with academic tasks (e.g., reading aloud, writing) and/or speaking with a teacher or peer about the assigned material. **Passive:** student is attending to (e.g., looking at, listening to) the assigned work (e.g. independent seatwork, teacher instructions, class-wide activities, group work). AEB does not include calling out or aimlessly looking around the classroom.

<table>
<thead>
<tr>
<th>Class: ________________</th>
<th>Observer: ____________________________</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date: _________________</td>
<td>Mon  Tue  Wed  Thurs  Fri</td>
</tr>
<tr>
<td>Phase: baseline</td>
<td>intervention  withdrawal  intervention 2  maintenance</td>
</tr>
<tr>
<td>Observation of: class-wide behavior</td>
<td>treatment integrity</td>
</tr>
<tr>
<td>Start Time: _______</td>
<td>End Time: _______</td>
</tr>
</tbody>
</table>

Description of Instruction (e.g., teacher directed instruction, small groups):

Notes:
Appendix D
Usage Rating Profile-Intervention (Revised)

Teacher Name: ___________________________ Date: ____________

Directions: Consider the Tootling* intervention when answering the following statements. Circle the number that best reflects your agreement with the statement, using the scale provided below.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tootling is an effective choice for addressing a variety of problems.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>I needed additional resources to carry out this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>I was able to allocate my time to implement this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>I understood how to use this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>A positive home-school relationship was needed to implement this intervention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>I am knowledgeable about Tootling procedures.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Tootling was a fair way to handle a child’s behavior problem.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>The total time required to implement the Tootling procedures was manageable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>I was not interested in implementing Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>My administrator was supportive of my use of Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>I have positive attitudes about implementing Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>Tootling was a good way to handle the child’s behavior problem.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>13</td>
<td>Preparation of materials needed for Tootling was minimal.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>Use of Tootling was consistent with the mission of my school.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>Parental collaboration was required in order to use Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Implementation of Tootling was well matched to what is expected in my job.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>17</td>
<td>Material resources needed for Tootling were reasonable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>18</td>
<td>I implemented Tootling with a good deal of enthusiasm.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>19</td>
<td>Tootling was too complex to carry out accurately.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>20</td>
<td>Tootling procedures were consistent with the ways things are done in my system.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>21</td>
<td>Tootling was not disruptive to the class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>22</td>
<td>I was committed to carrying out Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>23</td>
<td>Tootling procedures easily fit within my current practices.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>24</td>
<td>I needed consultative support to implement Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>25</td>
<td>I understood the Tootling procedures.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>26</td>
<td>My work environment was conducive to implementing Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>27</td>
<td>The amount of time required for record keeping was reasonable.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>28</td>
<td>Regular home-school communication was needed to implement Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>29</td>
<td>I required additional professional development in order to implement Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

* The term “Tootling” will be replaced with the name selected by the students in that class.
Appendix E

Children’s Usage Rating Profile

Date: ____________

Directions: Think about the Tootling* game you played in your class. After reading each sentence, circle the number that matches your belief about it. For example, if the sentence was “I like chocolate ice cream,” you might circle “4” for “I totally agree.”

<table>
<thead>
<tr>
<th></th>
<th>I totally disagree</th>
<th>I kind of disagree</th>
<th>I kind of agree</th>
<th>I totally agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tootling was too much work for me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. I understand why my teacher picked Tootling to help the class.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. I could see the class using Tootling again.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. Tootling was a good way to help students.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. It was clear what I had to do.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. I would not want to try Tootling again.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. Tootling took too long.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. If another class was having trouble, I would tell them to try Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9. I was able to do every step of Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10. I felt like I had to use Tootling too often.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11. Using Tootling gave me less free time.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12. There are too many steps to remember.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13. Using Tootling got in the way of other things.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14. I understand why the problem needed to be fixed.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15. Tootling focused too much attention on me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16. I was excited about Tootling.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Statement</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>---</td>
<td>---------------------------------------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>17.</td>
<td>Tootling made it hard for students to work.</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>18.</td>
<td>I would volunteer to use Tootling again.</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>19.</td>
<td>It was clear what the teacher needed to do.</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>20.</td>
<td>I was able to use Tootling correctly.</td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

* The term “Tootling” will be replaced with the name selected by the students in that class.
Appendix F

Tootling Intervention-- Teacher Script for Training Session

1. Define tootling.
   “Today we are going to talk about a new game that we’re going to play. We’re going to have a competition to see how many positive comments you all can make about each other. It will be a competition to earn rewards.”

2. Give examples of appropriate “positive comments.”
   “When we write down a positive comment, we focus on specific behaviors that we have seen with our own eyes that were appropriate. Behaviors we might see others doing that are appropriate are following the rules and being kind to others. Positive comments are NOT compliments about a person or something they have. A positive comment is saying that someone did something that was good.”

   Provide 2-3 examples of specific rule following behaviors and prosocial behaviors.
   Ex. “Jamar held the door open for Sally.”
   Ex. “Tatiana raised her hand and waited patiently to be called on.”

   Provide a non-example of an incorrect tootle about something a student has.
   Ex. “Bobby has a really cool new pen.”

3. Go over the procedure.
   “Every day you will be reminded to notice positive things you see your peers doing. On your desk will be these little papers (show paper). During class, if you notice one of your classmates doing something, you can write it down. At the end of class, I’ll collect the comments in this box (show box). Every day I’ll choose some to share with the class.”

4. Ask students to help name the game. Vote on new name.
   “Before we take a chance to practice writing down these positive comments, we need to come up with a name for this game. Anyone have ideas?”

   Some names that have been used with other classes include: Brags, Compliments, Kudos, and Positive Comments. Have students vote on the name.

5. Provide an opportunity for students to practice Tootling.
   “Now that you know how to make these positive comments, I’m going to hand out a paper and try to think of something you saw your classmates do today that was positive behavior.”

   Solicit volunteers to share out their positive comments. Or collect and read a few. Respond with praise or correction as students respond.

6. Tell the students they will be rewarded for tootling.
“Remember, this is a competition and with any competition there is a reward. Based on the reward, I will set a goal for the number of positive comments that the class needs to make in order to earn the reward. At the end of class each day, I will display the number of positive comments (show feedback chart) and we will discuss the class progress towards our class goal. If you have X number of positive comments, the class will earn a reward. I have some reward ideas, but I would like to get some ideas from you too.” Share the ideas and solicit ideas from students. Eliminate ideas that are not acceptable or possible in your classroom. Have students vote to select a reward they would like to earn.
1. **Remind students to be on the lookout for positive student behaviors.**
   “Remember what we said about “Positive Comments” (REPLACE WITH YOUR CLASS NAME) the other day. You are all in a competition to make as many positive comments about each other’s behavior as possible so that you can earn your reward.”

2. **Review procedures.**
   “When you see another student in class doing something good during class, write those behaviors on the cards I’ve just passed out. Put them in the box at the end of the period.”

3. **Show students the thermometer progress chart.**
   “This chart shows how many positive comments you’ve written so far. When you reach ____ positive comments, you will earn the class reward of __________.”

4. **Collect positive comments five minutes before the end of class and count up the total number of correct positive comments.**
   “Anyone have any other positive comments to add to the box?”

5. **Read aloud 5 of the positive peer reports.**
   “Here are a few of the positive comments from the day.” Read five comments.

6. **Praise the students for behaving appropriately, which earned them a positive comment.**
   “Nice job Bobby at picking up Suzy’s pencil when she dropped it.”

7. **Count up total number of correct positive comments. Disregard any jokes or inappropriate comments. Discuss the total number of positive comments and progress towards the goal.**
   “The class collectively wrote 25 positive comments today, your goal is 80 and now you only need to write 55 more to earn your reward.”

8. **Adjust the progress monitoring thermometer to display the total number of positive comments. Tell students the total number of positive comments they made and subtract this number from their overall goal.**

9. **If students met the goal for the reward, provide the reward to all students.**
   “Congratulations! The class needed to write ____ positive comments to earn the _______. Today you won! You wrote _____ positive comments and have earned _____.


Appendix G

Tootling Intervention Procedural Integrity Checklist

Observer Form—Teacher Training Session

Teacher: _____________________           Date: _______      Observer:_________________

Instructions: Mark an “X” in the Yes column if the primary investigator completed the component or an “X” in the No column if the primary investigator did not complete the component.

<table>
<thead>
<tr>
<th>Component</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Provided an overview of the Tootling intervention.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Defined “Positive Comments”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Gave examples of appropriate positive comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Gave non-examples of positive comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Provided and reviewed teacher with scripts for student training and intervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Provided and reviewed teacher with scripts for intervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Provided positive comments box</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Provided visual thermometer for recording cumulative positive comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Developed a list of possible rewards with teacher</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

<table>
<thead>
<tr>
<th>Number of Steps Completed (___/10)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Integrity %</td>
<td></td>
</tr>
</tbody>
</table>
Appendix H

Tootling Intervention Procedural Integrity Checklist

Observer Form—Initial Training Session with Students

Teacher: _____________________           Date: _______      Observer:_________________

Instructions: Mark an “X” in the Yes column if the teacher completed the component or an “X” in the No column if the teacher did not complete the component.

<table>
<thead>
<tr>
<th>Component</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Defined “Positive Comments”.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Gave examples of appropriate positive comments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Gave non-examples of positive comments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Solicited student ideas for names of competition. Voted on name.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Provided opportunity for students to practice writing positive comments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Provided feedback on at least 3 practice positive comments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Reviewed positive comments procedure.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Informed the class of the rewards for positive comments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Solicited student input into additional rewards.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Students voted on a reward.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

<table>
<thead>
<tr>
<th>Number of Steps Completed (___/10)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Integrity %</td>
<td></td>
</tr>
</tbody>
</table>
Appendix I

Tootling Intervention Procedural Integrity Checklist—Intervention Phase

*Observer Form*

Teacher: _____________________   Week: _______   Observer:_________________

Instructions: If the component is present, write “X”; if the component is not present, write “0.”

<table>
<thead>
<tr>
<th>Component</th>
<th>Date</th>
<th>Component Integrity Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feedback chart hung in a visible location in classroom</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>2. Positive Comment collection box visible</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>3. Paper slips for Positive Comments visible on students’ desks</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>4. Remind students to be on the lookout for appropriate peer behaviors</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>5. Review positive comments procedures</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>6. Discuss the cumulative Positive Comments and progress towards the goal</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>7. Read at least 5 Positive Comments at the end of the period</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>8. Praise the behaviors that earned Positive Comments</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>9. Sum the total number of Positive Comments and record on feedback chart</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>10. Reward the class when they meet the goal. (record N/A if class did not earn goal)</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Number of Positive Comments recorded</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session Integrity Percentage</th>
<th>/10</th>
<th>/10</th>
<th>/10</th>
<th>/10</th>
<th>/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix J

Tootling Intervention Procedural Integrity Checklist—Withdrawal Phase

Observer Form

Teacher: _____________________       Week: _______       Observer: ___________________

Instructions: If the component is present, write “X”; if the component is not present, write “0.”

<table>
<thead>
<tr>
<th>Component</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Component Integrity Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feedback chart not visible in classroom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/5</td>
</tr>
<tr>
<td>2. Positive Comment collection box removed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/5</td>
</tr>
<tr>
<td>3. Paper slips for Positive Comments not distributed/available.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/5</td>
</tr>
<tr>
<td>4. Teacher does not discuss recognizing positive peer behavior or positive comments procedures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/5</td>
</tr>
<tr>
<td>5. No Positive Comments read aloud at the end of the period</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/5</td>
</tr>
<tr>
<td>6. No rewards given to students for recognition of peer positive behavior.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/5</td>
</tr>
<tr>
<td><strong>Session Integrity Percentage</strong></td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
<td><strong>Overall Mean</strong></td>
</tr>
</tbody>
</table>


Appendix K

Maintenance Procedural Integrity Checklist

*Observer Form*

Teacher: _____________________           Week: _______      Observer:_________________

Instructions: If the component is present, write “X”; if the component is not present, write “0.”

<table>
<thead>
<tr>
<th>Component</th>
<th>Date</th>
<th>Component Integrity Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Positive Comment collection box visible</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>2. Paper slips for Positive Comments visible on students’ desks</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>3. Remind students to be on the lookout for appropriate peer behaviors</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>4. Review positive comments procedures</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>5. Read at least 5 Positive Comments at the end of the period</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>6. Praise the behaviors that earned Positive Comments</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Number of Positive Comments recorded</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Session Integrity Percentage | /6 | /6 | /6 | /6 | /6 |
| Overall Mean | | | | | |
Appendix I

Tootling Intervention Procedural Integrity Checklist—Intervention Phase

*Teacher Self-Report*

Intervention Teacher: _____________________       Week: _____________

Instructions: If the component is present, write “X”; if the component is not present, write “0.”

<table>
<thead>
<tr>
<th>Component</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Component Integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review instructions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Remind students of the rewards</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Show feedback chart of accumulated positive comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Provide index cards to students for positive comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Prompt students at end of class to place comments in the collection box</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Read at least 5 positive comments at the end of the period</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Praise the students who engaged in the publicly shared positive comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Sum the number of positive comments for the day and record on feedback chart</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Reward the class when they meet the goal. (record N/A if class did not earn goal)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/9</td>
</tr>
</tbody>
</table>

**Session Integrity**

<table>
<thead>
<tr>
<th>Overall Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>/9</td>
</tr>
</tbody>
</table>
Appendix M

Tootling Intervention Procedural Integrity Checklist—Withdrawal Phase

*Teacher Self-Report*

Intervention Teacher: _____________________  Week: ______________

Instructions: If the component is present, write “X”; if the component is not present, write “0.”

<table>
<thead>
<tr>
<th>Component</th>
<th>Date</th>
<th>Component Integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback chart not visible in classroom</td>
<td>Monday</td>
<td>Tuesday</td>
</tr>
<tr>
<td>Positive Comment collection box removed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper slips for Positive Comments not distributed/available.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher does not discuss recognizing positive peer behavior or positive comments procedures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Positive Comments read aloud at the end of the period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No rewards given to students for recognition of peer positive behavior.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session Integrity</td>
<td>/6</td>
<td>/6</td>
</tr>
<tr>
<td>Overall Mean</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix N

Tootling Intervention Procedural Integrity Checklist—Maintenance Phase

*Teacher Self-Report*

Intervention Teacher: _____________________   Week: ____________

Instructions: If the component is present, write “X”; if the component is not present, write “0.”

<table>
<thead>
<tr>
<th>Component</th>
<th>Date</th>
<th>Component Integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review instructions</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>Provide index cards to students for positive</td>
<td></td>
<td>/5</td>
</tr>
<tr>
<td>comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompt students at end of class to place comments</td>
<td>/5</td>
<td></td>
</tr>
<tr>
<td>in the collection box</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read at least 5 positive comments at the end of the</td>
<td>/5</td>
<td></td>
</tr>
<tr>
<td>period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Praise the students who engaged in the publicly</td>
<td>/5</td>
<td></td>
</tr>
<tr>
<td>shared positive comments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Session Integrity | /5 | /5 | /5 | /5 | /5 | Overall Mean |
