

2003

Evaluation of a whole-class token economy to manage disruptive behavior in preschool classrooms

Holly A. Filcheck
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Filcheck, Holly A., "Evaluation of a whole-class token economy to manage disruptive behavior in preschool classrooms" (2003). *Graduate Theses, Dissertations, and Problem Reports*. 1865.
<https://researchrepository.wvu.edu/etd/1865>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Evaluation of a Whole-Class Token Economy to
Manage Disruptive Behavior in Preschool Classrooms

Holly A. Filcheck

Dissertation submitted to the
Eberly College of Arts and Sciences
at
West Virginia University
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Child Clinical Psychology

Cheryl B. McNeil, Ph.D., Chair
Christina D. Adams, Ph.D.
Cindy M. Anderson, Ph.D.
Lindsey L. Cohen, Ph.D.
Tandy J. McClung, Ph.D.

Department of Psychology

Morgantown, West Virginia
2003

Keywords: Level System, Intrinsic Motivation, Behavior Problems,
Token Economy, Preschool, Disruptive Behavior

Abstract

Evaluation of a Whole-Class Token Economy to Manage Disruptive Behavior in Preschool Classrooms

Holly A. Filcheck

The behavior of children in a preschool classroom was assessed to evaluate the effectiveness of two classroom management approaches: (a) strategies already employed by the teachers, and (b) the Level System. Strategies already employed by the teachers were those that they currently used in their preschool classroom. This phase was considered the baseline or “A” treatment phase. The Level System is a new program that utilizes strategies including a token economy, stimulating rewards, strategic attention, and labeled praise to manage a range of behaviors exhibited by children in the classroom. This was considered the “B” treatment condition. These two approaches were evaluated using an ABAB sequence where each was employed for a minimum of 17 observations with 4 male 4-year-old participants. Behaviors were assessed using unobtrusive classroom coding of activities and teacher report. Four main findings were obtained: (a) the Level System was more effective in managing disruptive behavior than previously-used classroom management strategies, (b) fewer time outs were given while the Level System was used in the class compared to while the typical classroom management strategies were used, (c) teacher report of satisfaction with the Level System varied, but parents reported high levels of satisfaction with the Level System, and (d) negative effects on “intrinsic motivation” with use of the Level System were not evident.

Acknowledgments

This project could not have been completed without the unconditional support and guidance of Cheryl McNeil, Ph.D., committee chair, whose positive supervisory style has made this a wonderful experience. I would like to express my deepest gratitude and admiration to her. I also would like to sincerely thank my committee members, Christina Adams, Ph.D., Cindy Anderson, Ph.D., Lindsey Cohen, Ph.D., and Tandy McClung, Ph.D. for all of their excellent recommendations and assistance in making this a positive learning experience.

In addition, successful completion of this project also depended on my remarkable research assistants, Alyson Crumrine and Sara Simpson, who spent many long hours in the preschool assisting me with this project.

I also want to thank Kathy Zimmerman, director of the Morgantown Early Learning Facility (ELF), and the teachers who participated in this project because they permitted this study to be completed in their classrooms and dedicated many hours to training, completion of questionnaires, and consultation. Without their assistance, this project would not have been possible.

I am sincerely grateful to my parents, Patricia and William Filcheck, for their unconditional assistance and support throughout the completion of this project. Their positive support and love has made the accomplishment of my goals seem like an easy task.

Table of Contents

Title Page	i
Abstract.....	ii
Acknowledgments.....	iii
Table of Contents.....	iv
List of Tables and Figures.....	vi
Introduction.....	1
Token Economies.....	2
Considerations in Using Token Economies in Preschool Classrooms	5
Effectiveness of Token Economies.....	13
Using Token Economies to Manage Disruptive Behavior in Preschool Classrooms.....	18
The Level System	21
Summary.....	25
Purpose of Current Study.....	25
Hypotheses.....	25
Method	28
Setting.....	28
Participant Selection	29
Participants.....	29
Measures	30
Participant Selection Measures	30
Effectiveness Measures.....	31

Satisfaction Measures	32
Treatment Integrity Measures	33
Procedure	35
Teacher Training.....	35
Classroom Observations	36
Experimental Conditions	36
Criteria for Discontinuation of Data Collection.....	38
Results.....	38
Effectiveness.....	38
Satisfaction.....	47
Negative Effects.....	49
Treatment Integrity	49
Discussion.....	50
Main Findings.....	51
Methodological Limitations.....	55
Clinical Implications.....	59
Directions for Future Research.....	61
References.....	63
Appendices.....	73
Tables	97
Figures.....	102

List of Tables and Figures

Tables

Table 1. T-Scores of participants on CGI by teacher per condition	97
Table 2. Class management ratings by teacher per condition.....	98
Table 3. Mean number of time outs provided to participants by teacher per condition	99
Table 4. Percentage of integrity behavior exhibited by teachers in response to child behavior by condition	100

Figures

Figure 1. Percentage of intervals scored with appropriate behavior exhibited by each participant across teachers with horizontal lines indicating means for each phase	102
Figure 2. Percentage of intervals scored with appropriate behavior exhibited by Luke in each teacher's class with horizontal lines indicating means for each phase	103
Figure 3. Percentage of intervals scored with appropriate behavior exhibited by Cody in each teacher's class with horizontal lines indicating means for each phase	104
Figure 4. Percentage of intervals scored with appropriate behavior exhibited by Tim in each teacher's class with horizontal lines indicating means for each phase	105
Figure 5. Percentage of intervals scored with appropriate behavior exhibited by John in each teacher's class with horizontal lines indicating means for each phase	106
Figure 6. Mean percentage of intervals scored with appropriate behavior exhibited by each participant across teachers	107
Figure 7. Mean percentage of intervals scored with appropriate behavior exhibited by each participant in each teacher's class.....	108

Evaluation of a Whole-Class Token Economy to Manage Disruptive Behavior in Preschool Classrooms

Research suggests that the amount of behavior problems exhibited by preschool children (ages 2-5 years) has been increasing, and current estimates range from 2% to 17% (Campbell, 1990; Lavigne et al., 1996, 1998). The prevalence of behavior problems has increased from 1.7% in 1979 (Clark & Cheyne, 1979) to at least 8.3% in 1998 (Lavigne et al., 1996, 1998). For example, oppositional defiant disorder (ODD) has been reported to be present in 16.8% of preschool children (Lavigne et al., 1996, 1998). The reason for these increases is unknown. However, social factors such as increased parental stress, unemployment, single parent status, and decreased social support partially may account for this trend (Bacharach & Baumeister, 2003; Dunifon, Kalil, & Danziger, 2003). Thus, it appears that the number of children with behavior problems in preschool classrooms will increase (i.e., approximately 2-3 per class), and teachers will be required to manage these disruptive behaviors. However, 75% of teachers reported that their education insufficiently prepared them to manage children with special needs (including children with disruptive behavior disorders), and 72% of teachers reported that they were insufficiently prepared to manage behavior problems in the classroom (Merrett & Wheldall, 1993). Despite these facts, 69% of teachers reported that discipline is the most important issue to consider when attempting to improve the education of children (Merrett & Wheldall). Thus, providing preschool teachers with strategies to manage disruptive behavior in the classroom effectively is essential to meet the teachers' needs as well as the needs of the children. Furthermore, early intervention with children exhibiting behavior problems have yielded clinically significant improvements in functioning (e.g., Hoffman & DuPaul, 2000; McConaughy, Kay, & Fitzgerald, 1999; McNeil, Capage, Bahl, & Blanc, 1999).

The current study examined the effectiveness of the Level System (McNeil & Filcheck, in press), a new and innovative whole-classroom approach for behavior management in preschools (see Appendix A). The program is similar to a token economy with a response cost, and it was designed to provide preschool teachers with skills and strategies to manage behavior problems in their classrooms. The Level System consists of seven levels: three sunny levels, three cloudy levels, and one neutral level. Each child is assigned a shape that is placed on the chart. The teacher provides social reinforcement (i.e., labeled praise) and moves the children's shapes up a level for appropriate behavior. For inappropriate behavior, the teacher provides a warning and moves the children's shapes down a level for continued inappropriate behavior. At specified times throughout the day, all of the children with shapes in the sunny area receive a reward (e.g., activity, sticker). In the current study, potential positive and negative effects of the Level System were evaluated, as well as teacher and parent satisfaction with the program.

This paper is organized in the following manner. First, general issues surrounding the use of token economies in preschool classrooms are discussed. Second, considerations in using token economies in preschool classrooms are presented. Third, the effectiveness of token economies is discussed. Fourth, a review of the literature regarding the use of token economies to manage disruptive behavior in preschool classrooms is conducted. Fifth, a description of the Level System is presented. Finally, the results of the current study are presented and discussed with respect to the hypotheses.

Token Economies

Behavior management strategies (e.g., positive reinforcement, time out, response cost) have proven effective in decreasing inappropriate behavior and increasing appropriate behavior exhibited by children (e.g., Donahoe & Palmer, 1994; Martin & Pear, 1996; Miltenberger, 2000).

However, few such techniques have been evaluated in the context of managing disruptive behavior in preschool classrooms (Baker, Stanish, & Fraser, 1972; McGoey & DuPaul, 2000; Wolfe, Boyd, & Wolfe, 1983). In elementary classrooms, though, token economies have been used to manage disruptive behavior (e.g., Anhalt, McNeil, & Bahl, 1998; Bahl, McNeil, Cleavenger, Blanc, & Bennett, 2000; Drege & Beare, 1991). A token economy (e.g., star chart) has been defined as a program in which individuals earn tokens (e.g., poker chips, stickers) for exhibiting certain targeted behaviors (e.g., compliance). These tokens can be exchanged for rewards (e.g., activity, snack) (Cooper, Heron, & Heward, 1987; Martin & Pear; Miltenberger). Often, token economies include a response cost procedure as well. A response cost is characterized by the removal of a pre-determined number of tokens following the emission of an undesirable behavior (Cooper et al.; Martin & Pear). For the purpose of this paper, token economies will be discussed with the addition of the response cost procedure, unless otherwise specified.

A token economy is comprised of several components. First, the target behaviors must be operationally defined (Cooper et al., 1987; Miltenberger, 2000; O’Leary & O’Leary, 1977). These are the behaviors that the implementer wants to increase (e.g., compliance, sharing) or decrease (e.g., noncompliance, throwing toys). The operational definitions should be objective and explicit to the implementers as well as the children to ensure that each person understands which behaviors are required to receive tokens.

The second component involves determining the type of token, or secondary reinforcer, to be used (Cooper et al., 1987; Miltenberger, 2000; O’Leary & O’Leary, 1977). Miltenberger suggests using poker chips, smiley faces, coins, stamps, stickers, and geometric shapes as tokens, but warns against using tokens that can be duplicated or acquired elsewhere.

Another component of a token economy is establishing the backup or conditioned reinforcer (i.e., rewards) (Cooper et al., 1987; Miltenberger, 2000; O’Leary & O’Leary, 1977). The reward must be reinforcing to the child to increase the child’s appropriate behavior. In addition, in a classroom setting, the rewards also must be acceptable to the teacher. For example, some rewards may be logistically unacceptable to the teacher (e.g., long walks to fast food restaurants, expensive toys). Novel and variable rewards have been suggested to be effective with children, particularly those with behavior problems such as hyperactivity (Anhalt et al., 1998; Bahl et al., 2000). Additionally, variable reinforcers repeatedly have been shown to result in an increase in target behavior (e.g., Martin & Pear, 1996, Miltenberger).

Establishing a token exchange rate is the fourth component of developing a token economy (Cooper et al., 1987; Miltenberger, 2000). In other words, the implementer must decide how many tokens are required to receive each backup reinforcer. Typically, smaller rewards require fewer tokens than larger rewards. When using a token economy with preschool children, the exchange rates should be simplified to enhance understanding. Specifically, the amount of tokens needed to receive a reward should be kept small so that advanced counting skills are not required (e.g., 1 red sticker for a reward).

Deciding when and how often the tokens will be exchanged is the last component of a token economy (Cooper et al., 1987; Miltenberger, 2000). For example, a teacher may decide that tokens may be exchanged only at the end of the day, or after breakfast and after lunch. Depending on the reward, misbehavior may occur with the backup reinforcer. Specifically, children may want to play with the reward, or share it with other children. Rules concerning this behavior should be clear when the token economy is implemented.

The addition of a response cost to a token economy is suggested when there are

undesirable behaviors that potentially will compete with increasing the desired behaviors (Miltenberger, 2000). Adding a response cost procedure entails operationally defining the inappropriate behaviors that will result in the loss of tokens and the amount of tokens that will be lost for exhibiting the defined behaviors (Miltenberger; Walker, 1983). The number of tokens lost should reduce the likelihood that the child will have the opportunity to receive the backup reinforcer. However, the child must have some tokens in reserve or there will be no tokens to lose. Therefore, it is important to use baseline data to determine the amount of token loss per behavior (Cooper et al., 1987).

When determining whether or not to use a response cost in addition to a token economy, some disadvantages of response costs must be considered. For example, sometimes, the loss of tokens may be unethical such as removing food (Miltenberger, 2000). Also, using a response cost may result in increased verbal or physical aggression by the participant (Cooper et al., 1987). Other disadvantages of using response costs are that the implementer (e.g., teacher) may become a conditioned aversive stimulus leading to avoidance behavior from the student, and that the behavior being punished may actually be reinforced by the attention provided by removing the token (Cooper et al.).

Considerations in Using Token Economies in Preschool Classrooms

Practical considerations. Token economies have been shown to be easy to implement, acceptable to teachers and parents, and developmentally sensitive for preschool children (e.g., Jones, Downing, Latkowski, Ferree, & McMahon, 1992; McGoey & DuPaul, 2000; Storey, Danko, Ashworth, & Strain, 1994). However, token economies also may be expensive, overwhelming for teachers, and have weak maintenance and generalization (e.g., Corrigan, 1995; Herman & Tramontana, 1971; Miltenberger, 2000; Musgrove, 1981). According to Miltenberger,

the complexity of a token economy determines the amount of effort required in implementation and use. Each of the components of a token economy can have varying degrees of complexity. Therefore, in designing a token economy for a preschool classroom, the system should be simple to implement and use so that it does not interfere with the class activities, and so that the preschool children understand the token economy. For example, Storey et al. conducted a study in which the teachers' aides implemented a simple token economy with no response cost (i.e., praise and stickers given for social interactions) in order to increase social interactions during free play. According to reports by the teachers' aides, the token economy was easy to implement and use (Storey et al.), suggesting that simple token economies entail little effort from the teacher, which in turn, would disrupt the class activities less often than a complex token economy.

Research suggests that teachers, as well as parents, find token economies and response cost procedures to be highly acceptable (Elliott, Witt, Galvin, & Peterson, 1984; Jones, Eyberg, Adams, & Boggs, 1998; McGoey & DuPaul, 2000; Reynolds & Kelley, 1997; Witt, Elliot, & Martens, 1984). Specifically, Elliot et al. and Witt et al. conducted studies in which they used the Intervention Rating Profile (IRP) to determine teachers' acceptability ratings of behavioral procedures. In the study by Elliot et al., experienced teachers participated, whereas in the study by Witt et al. preservice teachers participated. The implemented procedures varied with respect to complexity (e.g., token economy is more complex than praise) and severity of child behavior problems (e.g., daydreaming to destruction of property). The authors found that teachers rated positive interventions (i.e., token economies, praise, and home-based reinforcement programs) as more acceptable than negative interventions (i.e., time out, response cost, and ignoring). In contrast, McGoey and DuPaul found that teachers rated a response cost procedure as more

acceptable than a token economy in their study. In addition, the teachers reported that the response cost procedure was easier to implement.

Jones et al. (1998) found similar results concerning token economies. More specifically, mothers of children with behavior problems rated 6 behavior management strategies using the Treatment Evaluation Inventory-Short Form (TEI-SF; Kelley, Heffer, Gresham, & Elliott, 1989). These strategies included: positive reinforcement, response cost, time out, differential attention, overcorrection, and spanking. Mothers rated positive reinforcement and response cost procedures as the most acceptable interventions compared to the others. Therefore, both time out and response cost procedures are acceptable to parents and teachers which is important because behavior management strategies that are more acceptable are more likely to be implemented efficaciously (Witt & Martens, 1983).

Some token economies have been shown to be developmentally sensitive so that preschool children can understand them (e.g., Jones et al., 1992; Swiezy, Matson, & Box, 1992; Titus et al., 1990). However, in order to be developmentally sensitive, there are several aspects of the token economy that must be considered. For example, token economies meet the developmental needs of preschool children when they are characterized by structure, predictability, simplicity, brightly-colored tokens, playfulness, visibility, and accommodations that are readily made for individual differences (Jones et al.; Kysela, 1972-1973; McGoey & DuPaul, 2000; Titus et al.). Exchange rates should be simple, and specific criteria for earning tokens should be set to facilitate understanding among preschool children. For example, research suggests that at about age 2 children begin to use numbers symbolically (Berger, 1998).

Swiezy et al. (1992) used a developmentally sensitive token economy with no response cost in their study. Specifically a bear puppet, “Buddy Bear,” was used to explain the

requirements of the token economy to children, and the tokens were brightly colored felt shapes that were age and gender appropriate (i.e., dinosaurs for boys, smiley faces for girls).

Other practical considerations include the financial cost of the rewards or backup reinforcers (Corrigan, 1995; Miltenberger, 2000), and the fact that teachers may have two to four children in their classroom with behavior problems. Given that the prevalence of behavior problems is increasing (Clark & Cheyne, 1979; Lavigne et al., 1996, 1998), teachers will continue to have more children that require behavior management programs. Therefore, teachers may have 2 to 4 token economies for different children in their class at one time. This may be overwhelming for the teacher because each token economy may use different tokens, different exchange rates, and different target behaviors.

Research suggests that token economies do not maintain or generalize reliably (e.g., Corrigan, 1995; Epstein, Masek, & Marshall, 1978; Herman & Tramontana, 1971; Miller & Schneider, 1970; Musgrove, 1981; Odom, Hoyson, Jamieson, & Strain, 1985). A few studies, however, suggested some generalization effects. For example, Swiezy et al. (1992) found that behavioral gains generalized across therapists but not across settings. In addition, behavior change was found to generalize to a non-intervention condition in a study by Miller, McCullough, and Ulman (1981). However, these results were obtained during a multi-element manipulation. Some authors suggest that using a fading procedure may aid in the maintenance of behavioral gains after the withdrawal of the token economy (e.g., Kysela, 1972-1973; Martin & Pear, 1996; Miltenberger, 2000; O'Leary, Poulos, & Devine, 1972; Storey et al., 1994). In addition, Corrigan suggests that implementers should foster realistic expectations in teachers concerning the lack of maintenance and generalization. Specifically, Corrigan stated that "no one expects that positive effects of psychopharmacological agents administered for a short time can

be maintained over time and across settings after the agents are withdrawn” (p. 1260). Thus, it may be unrealistic to expect lasting results from a brief token economy intervention after it is removed.

Increased occurrences of generalization are likely if the creator and implementer of the token economy first identify in which situations the target behavior occurs in order to promote generalization in that situation (Miltenberger, 2000). Also, identification of natural contingencies of reinforcement that will be available in the target behavior stimulus situation will promote generalization. The token economy should be designed to incorporate the stimulus situation and natural positive contingencies, and target behaviors should be monitored to determine if generalization is successful (Miltenberger). Continued assessment and manipulation of contingencies should continue until generalization is maintained.

Thus, token economies appear to be easy to implement and use in the classroom, acceptable to teachers and parents, and developmentally sensitive for preschool children. However, financial considerations, maintenance and generalization, as well as some philosophical issues should be considered with the use of token economies with preschool children.

Philosophical considerations. Some authors have suggested that philosophical issues are important when conducting research using token economies (e.g., Davidson & Bucher, 1978; Ford & Foster, 1976; Kohn, 1993, 2000; O’Leary et al., 1972; Turnbull, 1988). Specifically, philosophical issues concerning teachers, parents, and children may affect research in this area. However, these concerns, if addressed at the beginning of a study, may not result in complications throughout the study.

For example, authors of theoretical articles and chapters have suggested that because

token economies are straightforward (i.e., tokens are provided for specific target behaviors), the teachers' creativity and innovation with other behavior management skills may decrease (O'Leary et al., 1972; Turnbull, 1988). Furthermore, this decrease in creativity and innovation may extend to the curriculum in the classroom such that the teacher may become overly dependent on the token economy (Kohn, 2000; O'Leary et al.). For example, the teacher may use the token economy to hold the children's attention rather than developing a fun and exciting curriculum. However, supporters of using token economies (e.g., Miltenberger, 2000) continually recommend fading out the use of token economies so that natural contingencies begin to maintain target behavior. Therefore, the likelihood that teachers would become dependent on using a token economy would be minimized.

The possible decrease in "intrinsic motivation" that may result from using a token economy with preschool children has been researched extensively (e.g., Davidson & Bucher, 1978; Ford & Foster, 1976; Levine & Fasnacht, 1974; Kohn, 1993, 2000; Molloy, 1979; O'Leary et al., 1972; Roane, Fisher, & McDonough, 2003). For example, Lepper, Greene, and Nisbett (1973) conducted a study in which preschool children were rewarded for using specific drawing markers. Results suggested that they were less likely to use the markers after the rewards were withdrawn. In fact, the children used the markers less often than they did before the rewards were implemented. Thus, some authors have concluded that token economies should not be used with children for this reason. For example, Kohn (1993) argued that being rewarded only once for exhibiting a certain behavior can "kill your interest in it for weeks" (p. 74). Similarly, Levine and Fasnacht stated in a non-empirical article that token economies "should be avoided unless there is a real danger to the person or there is no alternative" because of the potential decrease in "intrinsic motivation" (p. 820).

Contradictory empirical evidence to this concern is abundant (e.g., Davidson & Bucher, 1978; Molloy, 1979; Roane et al., 2003). Results from this research suggested that token economies did not negatively affect intrinsic interest. Specifically, Molloy conducted a study in which 30 children were assigned to one of three conditions: token economy, expected reward (received reward without earning tokens), or unexpected reward. Tokens (which could be exchanged for rewards) or rewards (depending on condition) were given for drawing with colorful markers. The author found no significant differences between pre- and post-observations of drawing behavior for any of the conditions. In a study with an ABAB design, Davidson and Bucher provided 4 children with tokens for playing with a specific activity (i.e., house or clown). Results indicated that children did not engage in the reinforced activity less than they did during baseline once the reinforcement was withdrawn. Roane et al. conducted a study in which a 14-year-old male with delays was provided with a reward (20 s play with toy telephone or radio) for sorting silverware. Roane et al.'s findings suggested that when the participant was rewarded for sorting, his sorting behavior decreased, and when it was not reinforced, the behavior increased to above baseline levels. Thus, rewards did not decrease the participant's intrinsic interest in the task, but increased his interest.

Recently, Cameron, Banko, and Pierce (2001) conducted a meta-analysis using research over the past 30 years to determine the overall effects of rewards on "intrinsic motivation". They found that rewards produce no harmful effects during task performance. Specifically, results indicated that rewards produced positive effects on "intrinsic motivation" during low-interest tasks, and during high-interest tasks when they were explicitly tied to behavior and success. Negative effects on "intrinsic motivation" only were found when the rewards were expected, tangible, and not tied to the behavior. Thus, the authors concluded that rewards have no

pervasive negative effects on “intrinsic motivation”.

Kohn (1993, 2000) has suggested, in his books, that token economies create controlling environments that decrease children’s self-esteem because children exhibit certain behaviors (e.g., drawing with specific markers) only to receive the external rewards and not because they enjoy exhibiting that behavior. Kohn (1993) also has suggested that token economies promote competition among preschool students (i.e., children compete to receive the most tokens) which decreases teamwork and helping behaviors. Furthermore, this competition may lead to increased levels of anxiety in children because they are concerned with receiving tokens and rewards. Another concern raised by Kohn (1993) was that children may become dependent on receiving rewards and, therefore, may not engage in the targeted behaviors without expecting to receive tokens. According to Kohn (1993), this dependence is evident when the targeted behavior decreases after the withdrawal of the token economy. However, some authors (e.g., Reitman, 1998) have shown that Kohn (1993, 2000) ignored research (e.g., Dickinson, 1989; Vasta & Stirpe, 1979) that contradicted his viewpoints.

Concerns with fairness and the rewarding value of the classroom without the token economy are two other philosophical issues that have been discussed in conceptual works (Kohn, 1993; Skinner, Cashwell, & Dunn, 1996; Turnbull, 1998). Specifically, children with behavior problems may not exhibit targeted behaviors (e.g., sharing, compliance) frequently enough to receive the reward (Kohn; Skinner et al.). Therefore, they may not even attempt to receive tokens (Kohn). However, research has indicated that no disturbing effects have been found concerning the behavior of children who do not receive rewards (Okovita & Bucher, 1976). In addition, the token economy rewards may be more rewarding than the regular class activities, especially if the rewards are activity-based. Thus, the children may perceive the regular class activities as less

rewarding after the token economy has been implemented (Turnbull).

Several issues have been raised in the theoretical literature concerning parents. For example, if token economy charts are displayed in the classroom, the potential exists that other parents could have access to information regarding any child's behavior. Kohn (2000) has recommended that parents remove their children from classrooms in which behavioral charts (e.g., start charts) are displayed because it is evidence that students are being ranked and compared against each other. In addition, parents may not approve of using token economies if their children do not receive the rewards with enough frequency, and, as stated previously, children with behavior problems may not exhibit the targeted appropriate behaviors with enough frequency to earn the rewards (Kohn, 1993; Skinner et al., 1996). Similarly, Corrigan (1995) suggested that parents may feel that their children are being humiliated if they do not receive the rewards. Preliminary research (Filcheck et al., in press) in which parents responded to questions regarding these concerns suggests that most parents remain unconcerned about these issues.

In sum, several practical and philosophical issues exist with regard to using token economies in the preschool classroom. However, some research suggests that token economies may be a promising classroom intervention to manage disruptive preschool behavior (e.g., Baker et al., 1972; McGoey & DuPaul, 2000; Wolfe et al., 1983). Literature discussing the effectiveness of this intervention strategy follows.

Effectiveness of Token Economies

Little research has been conducted using token economies specifically to manage behavior problems in preschool classrooms (Baker et al., 1972; Filcheck, McNeil, Greco, & Bernard, in press; McGoey & DuPaul, 2000; Wolfe et al., 1983). However, research has examined the effectiveness of token economies with no response cost components in managing

preschool-aged children's disruptive behavior in non-classroom settings (e.g., home, hospitals) (e.g., Barkley, 1987; Budd, Leibowitz, Riner, Mindell, & Goldfarb, 1981; Herman & Tramontana, 1971; Murray, 1980; Musgrove, 1981; Swiezy et al., 1992). Also, the use of token economies to manage behavior problems in the classroom has been effective in decreasing disruptive behavior in elementary-aged children (e.g., Anhalt et al., 1998; Bahl et al., 2000; Drege & Beare, 1991).

Non-classroom settings. Herman and Tramontana (1971) conducted a study in which a token economy with no response cost was implemented in an experimental classroom with 6 preschool children with disruptive behavior. In addition, they attempted to generalize the resulting behavioral gains to a Head Start classroom. Group and individual token reinforcement procedures were used for appropriate rest-time behavior (i.e., staying on mat, making no verbalizations). In the experimental room, tokens were given for appropriate behavior during a game (i.e., throwing balls into a bin). Observations of appropriate rest-time behavior, then, were conducted in the classroom. Results suggested that when instructions regarding appropriate behavior for the game were given to the participants, they engaged in more appropriate behavior than when instructions were not given. In addition, differential effects concerning type of reinforcement (i.e., group or individual) were not found, and the results did not generalize to the regular classroom. Floor effects may have been a concern with this study causing any differences between group and individual contingencies to be undetected. Specifically, when instructions were given regarding appropriate behavior, the amount of inappropriate behavior decreased to near zero.

In a study by Rowbury, Baer, and Baer (1976), 7 "deviant" preschool children participated in two experiments, the "Baseline Experiment" and the "Guidance Experiment." In

these experiments, a mock preschool classroom was used with a teacher for experimental sessions. In the “Baseline Experiment,” the children received tokens for either one, two, or three task completions which could be exchanged for access to a play area. In the “Guidance Experiment,” the procedure was the same, differing only in the addition of analysis of teacher guidance (i.e., praise, prompts, instructions). Data obtained through behavioral observation indicated that the children’s task completion behavior increased significantly when teacher guidance was combined with the token economy. However, the effectiveness of each type of teacher guidance (e.g., praise) was not examined.

A study was conducted by Musgrove (1981) in which a token economy with no response cost was implemented in a facility for individuals with mental retardation with 3, 4-year-old preschool children who were emotionally disturbed. However, according to their full scale IQs (i.e., 69, 80, and 93), only one child exhibited mental retardation. During the token economy phase of the ABAB reversal design, children received stickers for targeted behaviors (i.e., staying seated and following commands). Through behavioral observation, results suggested that the token economy was effective in decreasing out-of-seat behavior and increasing compliance. However, these gains were not maintained when the children were transferred to public school. Musgrove suggested that this may have been because the reinforcers used in the school were not valuable to the children. Some concerns with this study are that the training and reliability of the observers was not reported, and the reversal period was only 1 day. Thus, stability was not reached during the reversal phase before the token economy was re-implemented.

Budd et al. (1981) conducted a study in which they taught parents to use a home-based reinforcement procedure. The children were able to earn tokens for absence of disruptive behavior (i.e., off-area, aggression, negative statements) in a summer treatment program. In

addition, in the home, parents reinforced their children's appropriate behavior. The methodology included a multiple-baseline design across behaviors for three groups of 6 children who exhibited disruptive behaviors. Behavioral observations were conducted, and results suggested that this delayed reinforcement procedure was effective. However, improvements were not found in the behavior of 2 of the 18 participants. Data were not collected regarding whether the parents actually gave the children rewards at home. Thus, these findings should be interpreted cautiously.

Swiezy et al. (1992) demonstrated that a token economy with no response cost could have a positive effect on preschool children with behavior problems. The Good Behavior Game, a token economy in which a puppet, "Buddy Bear," gives the children commands, was implemented with 4 children. The children were able to earn stickers for compliance to the given commands. The Good Behavior Game was implemented by therapists in a resource room or school kitchen separately for two dyads of children, and their free play behavior was randomly observed to assess for generalization affects. Results indicated that compliance and cooperation increased significantly during the treatment phase. Specifically, compliance rates increased from 11.7% and 27.3% during baseline to 74.7% and 76.5% during treatment. In addition, generalizability occurred across therapists but not settings. However, only two to four observation sessions during free play were conducted per dyad.

Elementary classroom settings. Drege and Beare (1991) used a multiple baseline design across students to examine the effects of a token economy with a response cost component and a time-out component as a backup consequence. Specifically, 3 male students in an emotional/behavioral disorder classroom participated. Children earned points for appropriate behavior (i.e., following rules, raising hand, using an inside voice) and lost points for

inappropriate behavior (i.e., breaking rules, speaking out in class, and lying). In addition, time out was used for severely disruptive behaviors (i.e., swearing, verbal aggression, aggression toward others, running away, and destroying property). Behavioral observation data indicated that the children's off-task and disruptive behavior significantly decreased during the treatment phase. However, the token economy and time out were not examined independently. Thus, the relative contributions of each could not be determined.

The use of a whole-classroom token economy has been shown to be effective in decreasing disruptive behavior in elementary classrooms (Anhalt et al., 1998; Bahl et al., 2000). A whole-classroom token economy is characterized by one token economy in which all of the children in the class (disruptive and typical) participate, and all receive the same reward (Anhalt et al.; Bahl et al.). In addition, most of the rewards are activity-based (Anhalt et al.; Skinner et al., 1996). Several practical and philosophical concerns (e.g., time, effort, financial constraints, children being singled out) mentioned earlier are addressed by using a whole-classroom token economy.

In Anhalt et al.'s (1998) study, a whole-classroom token economy with a response cost, The AD/HD Classroom Kit, also called The Tough Class Discipline Kit (KIT; McNeil, 1995; 2000), was implemented in a first grade classroom. Children were placed into groups (e.g., 4 or 5 children) and received labeled praise and happy faces for their group when exhibiting appropriate behavior and a warning and mild aversive consequences (i.e., sad faces) for their group when exhibiting inappropriate behavior. Rewards or backup reinforcers were given to groups with more happy faces than sad faces. Through behavioral observation of a target child (age 6) with disruptive behavior problems, results indicated that the child's level of appropriate and on-task behavior increased during the treatment as compared to the regular classroom

discipline program. In addition, the teacher and students rated the KIT as highly satisfying.

Bahl et al. (2000) used the same token economy and response cost program (i.e., KIT) as Anhalt et al. (1998). Bahl et al. implemented the KIT in two first-grade classrooms using a within-subject reversal design (i.e., ABA and BAB). Behavioral observations were conducted with 6 participants. Results suggested that rates of appropriate and on-task behavior were higher in the KIT condition as compared to the school-wide standard condition. Furthermore, teachers and students were highly satisfied with the use of the KIT, and teachers reported that their classrooms were less disruptive during the KIT condition.

In sum, token economies with time out components and whole-classroom token economies with response cost components have been proven effective in elementary classrooms (e.g., Anhalt et al., 1998; Bahl et al., 2000; Drege & Beare, 1991). In addition, token economies without response cost components have been shown to have positive effects on preschool children's behavior. Thus, these types of behavioral classroom management systems may be effective with preschool children. Although only one preliminary investigation of a whole-classroom token economy with preschool children was found in the present literature review (Filcheck et al., in press), a few studies have been conducted using individual token economies to manage disruptive behavior in preschool classrooms (Baker et al., 1972; McGoey & DuPaul, 2000; Wolfe et al., 1983).

Using Token Economies to Manage Disruptive Behavior in Preschool Classrooms

Research has demonstrated the effectiveness of token economies at managing disruptive behavior exhibited by children in preschool classrooms (Baker et al., 1972; Filcheck et al., in press; McGoey & DuPaul, 2000; Wolfe et al., 1983). Specifically, Baker et al. used an ABAB design in which a token economy with no response cost and a time-out procedure were

implemented to decrease the disruptive classroom behavior exhibited by 9 preschool children with mental retardation who were part of a larger class. The token economy consisted of praise and poker chips for appropriate behavior, and a 5-min time out for severe inappropriate behavior (e.g., tantrums and aggression). A control group of children in different classrooms who were matched for age and IQ were included. The poker chips could be exchanged for a tangible backup reinforcer (e.g., candy). The authors found that the experimental group engaged in significantly more disruptive behaviors at baseline, and significantly less disruptive behaviors during treatment than the control group. However, the number of time outs was not reported. Also, the authors did not assess the relative influences of time out and the token economy on treatment outcome.

Wolfe et al. (1983) implemented a token economy with no response cost with 3 preschool children with behavior problems in 2 settings (i.e., morning and afternoon classrooms) to examine the effects of a token economy on cooperative play. These children, who were a part of the larger class, received stickers placed on a “happy face chart” when they engaged in cooperative play (e.g., shared interaction between two children) for an entire min. The children could exchange a certain amount of stickers (i.e., eight or more) for 10 min of outside play. The authors found that when the token economy was in place, cooperative play increased by 50%, and time outs decreased compared to baseline. Furthermore, when the token economy was implemented in the morning, the children exhibited more cooperative behavior during the afternoon. Thus, generalization occurred between settings. This finding contradicts concerns that token economies decrease “intrinsic motivation.”

In a study conducted by McGoey and DuPaul (2000), a token economy and a response cost were implemented with 4 preschool children with Attention-Deficit Hyperactivity Disorder

(ADHD) who were a part of the larger class. A single subject reversal design (i.e., ABACABAC and ACABABAC) in two classrooms was used. During the token economy condition, the children could earn buttons for following classroom rules (e.g., keep hands and feet to self, stay in area), and five small buttons could be exchanged for one big button. Three big buttons were needed to receive the reward that was given at the end of the day (e.g., hand stamp, sticker). During the response cost condition, children lost buttons when they broke a classroom rule. In addition, a typically-behaved control child was observed as a peer comparison in each classroom. However, the peer comparison was observed less often than the target children. Results indicated, through direct observation and teacher ratings, that the children's disruptive behavior decreased during the token economy and response cost. Furthermore, teachers rated the response cost procedure as more acceptable than the token economy system. The authors did not report the data concerning the number of children that lost all of their tokens during the response cost phase or the average number of tokens lost.

One preliminary investigation was conducted with the Level System to compare the effectiveness and satisfaction of the system with Parent-Child Interaction Therapy (PCIT; see Hembree-Kigin & McNeil, 1995) skills (Filcheck et al., in press). Specifically, an ABACD treatment comparison design with a 4.5-month follow-up assessment was used in a preschool classroom referred for being "out of control." In this study, condition B was the Level System, and conditions C and D were the Child-Directed Interaction (CDI) and Parent-Directed Interaction (PDI) phases of PCIT respectively. Seventeen children (mean age = 2.9 years) and one teacher participated in the study. The teacher was trained in the use of CDI and PDI skills, as well as the Level System. Behavioral observations were conducted during a videotaped circle time in which the frequency of inappropriate behavior exhibited by any participant in the

classroom was coded. Thus, data were combined across children. Results suggested that the frequency of inappropriate behavior exhibited by the children decreased throughout the study while the number of time outs given by the teacher increased throughout the study. Teacher report indicated that she was more satisfied with using PCIT skills than the Level System. However, the teacher chose to use the Level System at the study's end. This preliminary investigation was limited by the lack of reversal of inappropriate behavior during the withdrawal phase, and the low level of teacher treatment integrity with the Level System. However, it provides preliminary support that the Level System may be a viable option for the management of disruptive behavior in preschool classrooms.

In sum, these studies (Baker et al., 1972; Filcheck et al., in press; McGoey & DuPaul, 2000; Wolfe et al., 1983) demonstrate that token economies and response costs are effective in decreasing disruptive behavior and increasing compliance and cooperation among preschool children with behavior problems. However, the current literature search revealed only one study (Filcheck et al.) that was conducted using whole-classroom token economies with response costs in preschool classrooms. The effectiveness of the Level System will be examined in the current study.

The Level System

Because of the effectiveness of individual token economies with preschool children with behavior problems (e.g., Baker et al., 1972; McGoey & DuPaul, 2000; Wolfe et al., 1983), and the effectiveness of a whole-class token economy with elementary school children (Anhalt et al., 1998; Bahl et al., 2000), it was hypothesized that a whole-class token economy for preschool children would be effective in managing minor disruptive classroom behavior (e.g., whining, yelling, noncompliance). Thus, the Level System (McNeil & Filcheck, in press) was developed

to address the issue of increased numbers of children with behavior problems in the preschool setting. The Level System incorporates several of the techniques (e.g., labeled praise, warning signal) that are involved in the KIT (see Anhalt et al.; Bahl et al.; McNeil, 1995; 2000).

However, the techniques have been made developmentally appropriate for preschool children. The Level System also was designed to address several practical and philosophical concerns regarding the use of token economies (i.e., ease in implementation, developmental sensitivity, small financial cost to teachers, fairness, displaying behavioral information).

The Level System is unique in that it is designed for the entire classroom. One advantage of using a whole-classroom approach is that no child is singled out. In other words, all of the children will have their behavioral information (i.e., tokens) displayed in the room, and have the opportunity to receive the rewards. This addresses several philosophical concerns discussed previously. In contrast, an individual token economy (e.g., star chart) only allows for one child's tokens to be displayed and for that child to receive rewards, which may be seen as unfair by other children or parents. In addition, the Level System meets the needs of the entire class because effective behavior management strategies (e.g., praise, tokens, response cost) are integral to the program (Cooper et al., 1987; Martin & Pear, 1996; Miltenberger, 2000).

The Level System consists of a chart with seven levels. The top three levels, the "sunny area," contain pictures of suns with smiling faces, and the bottom three levels, the "cloudy area," contain pictures of clouds. Between the two areas is a neutral level (see Appendix A). Each child is assigned a certain shape on the System (e.g., dinosaur, boat, train, heart), and that child's name is written on the shape, which is placed in the neutral area at the beginning of each reward period. Consequences are given for both appropriate and inappropriate behavior at least 10 times per hr. Specifically, the children's shapes are moved up a level for appropriate behavior and

down a level for inappropriate behavior. The children's shapes need to be placed at a specified level (i.e., any of the sunny levels) in order to receive a reward.

The Level System has been designed to address several practical concerns when considering use of token economies. Specifically, it was designed to be developmentally sensitive for preschool children and practical for teachers. For example, the Level System is developmentally sensitive in that children's shapes simply are moved up a level on a chart after engaging in appropriate behavior, and down a level after engaging in inappropriate behavior. Therefore, instead of earning tokens for appropriate behavior and exchanging them for rewards (which requires mathematic skills), all children that are in the sunny area on the chart at a particular time receive a reward. The behavioral goals of the Level System are developmentally sensitive in that the pre-established classroom rules (e.g., keeping hands and feet to self, staying in area, playing gently with toys, following instructions) will be the "appropriate" behaviors that will be reinforced to increase the compliance with these rules. The developmental sensitivity of the Level System extends to the classroom materials it utilizes as well. For example, the chart contains brightly-colored shapes (e.g., circle, airplane, truck) that are familiar to 2- to 5-year-old children.

In addition, this system has been developed to be practical for preschool teachers. For example, because the teacher simply moves the children's shapes on the chart, and all children receive the same reward, the teacher would not have to dispense tokens or different rewards to each child, which could be time consuming. Also, training in the Level System requires approximately one hr, whereas other methods of training to manage disruptive behavior may require more time and effort. Specifically, McIntosh, Rizza, and Bliss (2000) trained a preschool teacher in the use of Teacher-Child Interaction Therapy (TCIT), an intervention using positive

skills (e.g., praise, description), giving effective commands, time out, role plays, homework, and in vivo coaching of skills, to manage disruptive behavior. Twelve 1-hr sessions were required in order for the teacher to learn the entire program. Therefore, developing an intervention that requires much less training time and effectively manages disruptive behavior would benefit teachers as well as professionals consulting in the classroom.

As mentioned before, children's shapes are moved up for appropriate behavior and down for inappropriate behavior. In addition, when their shapes are moved up, they are given a labeled praise for the appropriate behavior (i.e., specific praise such as "Thank you for sitting in your seat") by the teacher. Also, children are given a warning for inappropriate behavior before their shapes are moved down a level. For example, the teacher would hold up two fingers and say, "You have two choices. You can either play gently with the blocks, or you will move down a level." If after the warning, the child does not begin to behave appropriately, the teacher moves the child's shape down a level. Teachers are instructed to use a monotonous tone while giving a warning to provide minimal attention to inappropriate behavior. Teachers are encouraged to give frequent and immediate feedback (i.e., enthusiastic labeled praise, warnings, moving shapes) for appropriate and inappropriate behavior.

At the end of a certain period of time (e.g., after circle time), all children with shapes in the sunny area of the Level System receive the same reward (e.g., snack, activity, sticker), and all children with shapes in the cloudy area do not receive the reward. The children who do not receive the reward will participate in the regular classroom activity (e.g., coloring, painting) while the other children participate in the reward. After the reward is given, the children who received the reward go back to the regular classroom activity. The rewards are printed on cards, and the teacher chooses a reward card from the stack to employ (see Appendix B). The teacher

must provide each reward on the cards once before reusing any reward. Most of the rewards are activity based (e.g., opportunity to play a game such as charades) to reduce cost of backup reinforcers to the teacher (Kysela, 1972-1973; Miltenberger, 2000) and address a practical concern. In addition, all of the children receive the same reward to reduce the time involved with providing the rewards. The rewards are distributed one to two times in the morning and one to two times in the afternoon because children with behavior problems need frequent feedback and consequences (i.e., positive and negative) to manage their behavior effectively (e.g., Barkley, 1987; Hembree-Kigin & McNeil, 1995).

After the reward is given to the children, all shapes are placed back in the neutral area, and a new period begins during which they can earn a reward. The children, essentially, are starting over for the next period in order to maintain their motivation level. In other words, if a child who was at the bottom of the cloudy zone did not receive “another chance” at acquiring a reward, he or she may not be motivated to behave appropriately.

Summary

In sum, research has indicated that token economies are an effective means of managing disruptive behavior exhibited by preschool children. In addition, whole-classroom token economies have been proven effective in elementary classrooms. However, only one study (Filcheck et al., in press) was found that examined the effectiveness of a whole-classroom token economy in managing behavior problems in preschool classrooms. Most of the studies conducted with this population thus far have used individual token economies for each child (e.g., star charts). A whole-class behavior management system seems essential with the increasing amount of preschool children with disruptive behavior problems and the need for teachers to have a simple, time-efficient and effective way to manage classroom behavior.

Purpose of Current Study

The purpose of the present study was to examine and compare the effectiveness of two classroom management strategies in managing disruptive behavior exhibited by preschool children: the typical classroom management strategy (i.e., baseline, no intervention) and the Level System (i.e., whole-classroom token economy). Specifically, the current study evaluated whether or not the children's appropriate behavior increased and inappropriate behavior decreased while the Level System was used in the classroom, as compared to the strategies already utilized by the teachers. The effectiveness of these classroom management approaches was determined by behavioral observation and teacher report of disruptive behavior. In addition, satisfaction with the intervention was assessed by teacher and parent report of satisfaction as well as which approach the teachers chose to continue after termination of the study. Lastly, possible negative effects on "intrinsic motivation" were examined by visual inspection of the data. For example, decreases in appropriate behavior to below baseline levels upon removal of the Level System would have suggested that the children were no longer "intrinsically motivated" to engage in appropriate behavior (see Davidson & Bucher, 1978; Ford & Foster, 1976; Levine & Fasnacht, 1976; Kohn, 1993, 2000; Molloy, 1979; O'Leary et al., 1972). Because the Level System is a whole-classroom approach, it is important to assess behavioral changes in children with varying levels of disruptive behavior. Therefore, children with behavior problems and children with typical behavior participated.

Hypothesis one: Effectiveness. It was hypothesized that the Level System would be more effective than the management strategies that the teachers were already using in promoting prosocial behavior in children with disruptive behavior as well as children with typical behavior. In other words, the children were expected to exhibit more appropriate behavior during the Level

System conditions, than during the baseline conditions. This result was expected because previous research suggests that individual token economies are effective in managing disruptive behavior exhibited by preschool children with disruptive and typical behavior (e.g., Baker et al., 1972; McGoey & DuPaul, 2000; Miller, McCullough, & Ulman, 1981; Okovita & Bucher, 1976; Wolfe et al., 1983) and whole-classroom token economies are effective in managing disruptive behavior exhibited by elementary children (i.e., disruptive and typical) (Anhalt et al., 1998; Bahl et al., 2000). In addition, a preliminary study indicated that the Level System was effective in managing challenging behaviors in a preschool classroom (Filcheck et al., in press). Furthermore, the Level System is based on behavioral principles that have been proven effective (Cooper et al., 1987; Martin & Pear, 1996; Miltenberger, 2000), and it has been designed specifically to manage disruptive behavior in preschool children.

Hypothesis two: Satisfaction. It also was expected that teacher satisfaction with the Level System would be associated with improvements in the children's behavior. Specifically, it was hypothesized that the teachers would be most satisfied with the behavior management approach that was more effective, and as stated above, it was expected that the Level System would be the more effective system. This result was expected because in previous studies, teachers have reported high satisfaction with token economies and response cost procedures (Elliott et al., 1984; McGoey & DuPaul, 2000; Reynolds & Kelley, 1997; Witt et al., 1984). In addition, teachers have reported high satisfaction with whole-classroom token economies because they require minimal additional time, effort, and resources to implement (Anhalt et al., 1998; Bahl et al., 2000; Filcheck et al., in press).

Hypothesis three: Negative effects. It was hypothesized that the token economy would produce no negative effect on "intrinsic motivation" for the participants. Specifically, when the

token economy program was removed from the classroom, it was expected that the children's level of appropriate behavior *initially* would decrease below baseline, but over time, it was expected to increase and remain stable at baseline levels. This finding was expected because research that examined the effects of token economies on preschool children's "intrinsic motivation" suggested that no lasting negative effects on intrinsic interest were evident (e.g., Davidson & Bucher, 1978; Molloy, 1979). Therefore, the same results were expected for a whole-classroom token economy.

Method

Setting

Data were collected from October through April in a rural preschool with one class and 5 preschool teachers. This preschool was arranged such that the class of children rotated through 2 teachers per day for 1 hr classes in the morning. The classes included: arithmetic, Spanish, computer, reading, and science. For example, the class may be with the computer teacher from 9:00 a.m. to 10:00 a.m., and then move to the reading teacher's room from 10:00 a.m. to 11:00 a.m. The participating class included 13 children ($M = 3.5$ years, $SD = .52$, range = 3-5 years). Most of these children were from homes in which their parents were married (85.2%), had professional jobs (74.1%), obtained either advanced (25.9%), master's (22.2%), or bachelor's (22.2%) degrees, and made over \$60,000 per year (55.6%). All conditions of the study were conducted in the regular classroom with the acting primary teacher (one of five) and teacher's aide who changed frequently (i.e., many different individuals were hired for this position). If the teacher was absent, the data for those observations were not included because the substitute teacher was not expected to use the behavior management approaches in the same manner that the teacher would use them (i.e., substitute teachers were not trained by the experimenter in the

use of the Level System). This occurred twice throughout the study. If, on the other hand, a child participant was absent, data collection continued for the participants who were present.

Participant Selection

Parental informed consent was obtained for 100% of the children in the class. Thus, they all were considered for participation in the study. Two preschool children with typical behavior and 2 preschool children with disruptive behavior were selected for participation based on teacher's report on the Conners' Global Index (CGI, Conners, 1997) (Appendix C). Even though all the children in the classroom received the intervention, only the behavior of these 4 children were assessed. Children with the high scores on the CGI (i.e., gender and age-corrected T-score ≥ 60) were considered children exhibiting disruptive behaviors, and children with scores that approximated the average of the class (i.e., gender and age-corrected T-score 45-55) were considered children exhibiting typical behavior. Three children met criteria for the disruptive behavior category. Therefore, the participants for this category were randomly chosen. Four children met criteria for the typical category. Participants in this category were chosen to match the race and gender of the children with disruptive behavior. Teachers were unaware of which children were participating in the study.

Participants

The 2 children with disruptive behavior who were selected for participation were "Luke" and "Cody." Both of these children were 4 year-old Caucasian males from two-parent homes. The children's scores on the CGI were averaged from all of the teacher's ratings. On the CGI, Luke received a T-score of 60.44 and Cody received a T-score of 70.44. Both of these scores were considered to be in the disruptive category. The 2 children with typical behavior who were selected for participation were "Tim" and "John." Both of these children were 4 year-old males

and from two-parent homes. Tim was Caucasian and received a T-score of 45.20 on the CGI, and John was Bi-racial and received a T-score of 45.00 on the CGI. Additionally, 5 teachers (“Patricia,” “Colleen,” “Lori,” “Chad,” and “Rachel”) participated in the study (4 female and 1 male). They ranged in age from 25 to 63 ($M = 40.8$, $SD = 15.6$), and all were Caucasian. All teachers possessed a bachelor’s degree, and Colleen and Chad were working to obtain master’s degrees. Amount of teaching experience ranged from 2 to 10 years ($M = 6.0$, $SD = 3.8$). The names of all children and teachers were changed for confidentiality purposes.

Measures

Participant Selection Measures

Conners’ Global Index (CGI). The CGI is a shortened 10-item version of Conners’ (1969) original scale that had 39 items. The CGI assesses disruptive behavior as well as hyperactivity. It contains disruptive behaviors rated on a 4-point Likert scale ranging from “not at all” (0) to “very much” (3). This index is sensitive to detecting behavioral changes due to treatment, and it requires a short amount of time to administer (Conners, 1997; Wainwright & MHS Staff, 1996). Several studies demonstrate the reliability, validity, and usefulness of this measure (e.g., Brown & Wynne, 1982; Diamond & Deane, 1990; DuPaul, 1991; DuPaul & Barkley, 1992; Epstein & Nieminen, 1983). Specifically, it demonstrates high inter-rater reliability (.85) and high internal consistency (.90) (Margalit, 1983). The teachers completed a CGI for every student in the class (Appendix C) at the beginning of the study as well as after each condition.

Demographic Questionnaire. This form is a short questionnaire that includes the child’s age, gender, race, marital status of caregivers, family employment, and education and income level of caregivers. It was used to determine whether demographic differences exist between the

children with typical behavior and the children with disruptive behavior. All parents were asked to complete a short demographic questionnaire (Appendix D).

Effectiveness Measures

Revised Edition of the School Observation Coding System (REDSOCS). The REDSOCS (Jacobs et al., 2000) is a behavioral observation coding system that is the revised version of McNeil, Eyberg, Eisenstadt, Newcomb, and Funderburk's (1991) School Observation Coding System (SOCS) which operationalized several classroom behaviors (e.g., appropriate vs. inappropriate, on-task vs. off-task, and compliance vs. noncompliance). However, in the current study, only the appropriate and inappropriate behavior categories of the REDSOCS were recorded (see Appendix E for definitions). This behavioral category was chosen because preschool children do not often have "tasks" that they must attend to because much of the class is unstructured. Even during circle time, the children are not required to be "on-task," only to behave appropriately. In addition, Bahl et al. (2000) found that little information was gained from the compliance category because many teacher commands are indirect or implicit. Thus, behaviors falling into these categories were not recorded. In addition, the definition of "cheating" within the inappropriate behavior category was not used in the current study because it is inapplicable to the preschool classroom.

The REDSOCS is designed to measure disruptive classroom behavior and uses a partial-interval coding system. Specifically, the REDSOCS uses a 10-s observation system in which behaviors are marked at the end of the interval with no pause except to rotate children at the end of each min. The recording procedure for the current study was a 10-s observe, 5-s record schedule conducted for one participant at a time for approximately 15 min during structured morning activities (Appendix F) with the order balanced across participants. Thus, the recording

procedure was modified. This behavioral observation system was used to code the behavior of the 4 participants in the current study. Jacobs et al. (2000) demonstrated that the REDSOCS has good psychometric properties. Specifically, inter-rater agreement of the coders for appropriate and inappropriate behaviors were .85 and .83 respectively.

Class Management Rating. The teachers were asked to complete a daily manageability assessment of the children's behavior. This rating was made on a 5-point Likert scale from "completely unmanageable" (1) to "completely manageable" (5) (see Appendix G). This assessment provided information about how manageable the class as a whole was during different conditions.

Time-Out Log. The time out log was completed daily by the teachers throughout the study. The log included which children received time outs each day, why they received time outs, and how long the time out lasted (Appendix H). Although time out was not addressed directly by the intervention in the current study, this information aided in assessing whether more time outs occurred in one condition as compared to the others, and thus, the effectiveness of the behavior management procedures (i.e., Level System and regular classroom strategies).

Satisfaction Measures

Intervention Rating Profile (IRP). The IRP (Witt & Martens, 1983) is a 20-item teacher report measure that assesses acceptability of interventions on a 6-point Likert scale from "strongly disagree" (1) to "strongly agree" (6) (Appendix I). The inventory results in one major factor as determined by factor analysis, overall acceptability of the intervention. Higher scores represent greater acceptability. Research suggests that the IRP is reliable (.91), and sensitive to differences concerning the acceptability of different treatments (Witt et al., 1984; Witt & Martens). The teachers completed this measure at the end of each condition and at the follow-up

assessment.

Parent interview. The parental interview is a 15-item form that was constructed for this study to assess parental views concerning classroom management strategies (e.g., redirection, token economies, the Level System). The first 10 items are rated on a 6-point Likert scale from “very unacceptable” (1) to “very acceptable” (6), and the last 5 items are in an open-ended format (Appendix J). Parents completed this form once, at the end of the study.

Treatment Integrity Measures

Training coders. Two advanced undergraduates were trained as research assistants to complete the classroom behavioral observations using pilot training videotapes and definitions (Appendix E, F, and K) of child appropriate vs. inappropriate behaviors as well as teacher behaviors (see Appendix K). Teacher behaviors included: praise, warning signal, providing a child with time out, moving shapes up a level, moving shapes down a level, criticism, and other behavior. A time out was coded when a teacher indicated placing a child in “time out.” “Other” included interactions between the teacher and the child that did not meet the requirements of the other codes (e.g., reading a book, sitting on the teachers lap, engaging in an activity). Training occurred until the research assistants independently obtained 80% agreement for three consecutive training sessions with the author on videotaped observations of preschool classrooms for all behaviors. Then, the coders conducted live observational recording in the classroom independently.

Interobserver agreement. Agreement was assessed throughout data collection for 30-35% of the observations on the following variables: inappropriate behavior, praise, moving shapes up a level, moving shapes down a level, time out, warning, criticism, and other. Agreement observations were chosen randomly throughout data collection. A dual headphone jack was used

and the observers had as much distance as possible between them. If on any of these observations, agreement fell below .75 Kappa, which is considered excellent (Fleiss, 1981), or 80% mean occurrence/nonoccurrence agreement, for 2 consecutive observations, the research assistants were retrained before independent coding continued. The Kappa statistic fell below .75 for 2 consecutive observations 2 times for the “other” category of teacher behavior (range .211 - .828). According to the mean occurrence/nonoccurrence agreement, this occurred 2 times (range 25.3% - 86.1%) as well. Therefore, retraining of coders occurred twice throughout the study. The mean Kappa for each behavioral category was as follows: inappropriate behavior = .83, praise = .83, criticism = .80, moving shapes up a level = .96, moving shapes down a level = .98, time out = 1.0, warning = .88, and other = .77. Additionally, the mean occurrence/nonoccurrence agreement for each behavioral category was as follows: inappropriate behavior = 84.8%, praise = 84.8%, criticism = 84.4%, moving shapes up a level = 96.2%, moving shapes down a level = 97.9%, time out = 100%, warning = 88.7%, and other = 78%. The research assistants were unaware of the participants’ assignment and the study’s hypotheses.

Treatment integrity measures. Treatment integrity measures included behavioral observation of the teachers’ responses to the children’s behaviors (e.g., labeled praise, moving child’s name up a level, warning, moving child’s name down a level, time out, criticism, other) (Appendix F), as well as the daily completion during the treatment condition of an integrity checklist by the coders (Appendix L). These measures assessed the accuracy of the teachers’ implementation of the Level System. The teachers were informed that these measures were completed. A score of 85% or greater on the treatment integrity checklist was considered an accurate and successful implementation of the treatment for that observation. If treatment integrity on the checklist was less than 85% for 2 consecutive observations, the data for those

particular observations would not have been included and the teacher would have been retrained in the procedure of the Level System. Data collection would not have begun again until integrity reached 85%. However, this criterion was never met throughout the study (range 0%-100%; $M = 92.1$). Integrity was below 85% on 2 observations, but they were not consecutive (58.3% and 0%). Thus, retraining did not occur. Integrity reached 0% when a teacher forgot to use the Level System. Data for the observations with questionable integrity were not used.

Additionally, the teachers were asked to construct a list of which children received and did not receive rewards to ensure that most children received the rewards on most days (Appendix M). Specifically, this measure indicated that approximately 2 children did not receive rewards during each observation. However, the children not receiving the rewards changed. On one occasion, on observations 5, 6, and 7 during the first Level System Condition, this measure indicated that a target participant with disruptive behavior did not receive the reward. Thus, it was discussed with the teachers that the individual expectations for this child be changed so that he had the opportunity to receive the reward. At no other time during the study did this occur.

Procedure

Teacher Training

The teachers were trained in the use of the Level System 3 days before the baseline condition ended. Teacher behavior on these 3 days was monitored to ensure no changes in teacher behavior occurred before implementation of the Level System (Condition B). Visual inspection of teacher praise and criticism and child appropriate behavior indicated that no differences were evident. The experimenter and Cheryl McNeil, Ph.D., provided a 2-hr workshop to the teachers to explain the use of the Level System. In addition, the teachers were trained to use labeled praise and avoid criticism (see Appendix J) which are integral parts of the

Level System (i.e., social reinforcement). At the beginning of the Level System condition, the experimenter provided in-vivo coaching to the teachers in the use of the Level System until 85% treatment integrity was reached. Integrity was reached for each teacher after one coaching session. Finally, the experimenter provided feedback to the teachers throughout data collection regarding his or her treatment integrity. Specifically, after each observation, treatment integrity checklists were reviewed, and the teachers were provided with verbal feedback regarding their implementation once a week.

Classroom Observations

Classroom observations were conducted separately by two research assistants approximately four to five times each week, in the morning, except for 30-35% of all observations which were conducted jointly by a research assistant and the experimenter to assess inter-rater agreement. These observations were conducted during one class period per day in the morning for approximately 1 hr. As stated earlier, the REDSOCS was used to record the exhibited behaviors on a 10-s observe, 5-s record partial-interval system conducted for 1 participant at a time for 15 min each with the order balanced across participants. Additionally, the research assistant recorded the teachers' responses to the child's behavior (e.g., labeled praise, moving child's name up a level, warning, moving child's name down a level, time out, criticism, other) using the same observe/record partial-interval system during the same observation periods. In other words, the child and teacher behaviors were coded concurrently. The coders used an audiotape with earphones to cue them to the child to be coded and when observations and recordings were to occur.

Experimental Conditions

The proposed study used a single subject withdrawal design (i.e., ABAB) with a 1-month

follow-up assessment. Two conditions were included in this design: the classroom management strategies that the teachers already used (“A”), and the Level System (“B”). The teachers were asked to use strategies already in place and the Level System until stability was reached or at least 15 observations had passed, whichever was shorter. Stability was defined as at least 8 observations of data collection as well as no clear trends in any child’s appropriate behavior for the last 3 consecutive observations of data collection. Even though some participant’s data appeared to be stable in some conditions (e.g., see baseline for Cody in Figure 1), all participants’ data was not stable for any condition. Therefore, the stability criterion was never reached in any condition, thus, each condition was maintained until each child was observed for at least 15 observations. Conditions were held for an increased amount of time to ensure that any changes in data were the result of treatment. Condition A was held for 19 observations. Each of the Level System conditions was held for the following number of observations respectively: 17, 20. After the withdrawal of the Level System (i.e., withdrawal phase), condition “A” was to be in place until stability was reached, plus 1 week, to examine possible negative effects on “intrinsic motivation.” However, because stability was not reached, this condition remained implemented until each child was observed for 15 observations, plus 1 week, totaling 24 observations. The teacher used whichever management strategies were most satisfying after the experimental conditions ended. Thus, a 1-month follow-up assessment was conducted to determine which intervention was being used and the amount of disruptive behaviors exhibited by the participants. The follow-up assessment was to last at least as long as the shortest condition. Therefore, this condition was conducted for 17 observations. If the teachers were using the Level System at the 1-month follow-up assessment, treatment integrity was to be assessed. However, none of the 5 teachers was using the level system during follow up.

Strategies already used in the classroom. The teachers were asked to use the techniques that they already used to manage classroom behavior during the first and third conditions of the study. These conditions served as the baseline and withdrawal phases, respectively. Strategies utilized by the teachers included: verbal reprimands, yelling, redirection, time out, and removal from the class.

The Level System. The teachers were asked to use the Level System, with integrity, during the second and fourth conditions. These conditions served as the treatment. The teachers were required to obtain and maintain 85% integrity using the Level System. If integrity was less than 85% for 2 consecutive observations, retraining would occur. However, the teachers provided the intervention with acceptable levels of integrity, and retraining never occurred.

Criteria for Discontinuation of Data Collection

To protect against compromising the integrity of the study, specific criteria, if met, would have resulted in the discontinuation of data collection and beginning the study in a new classroom. Specifically, the study would have begun in a new classroom if: (a) there were ceiling effects (i.e., the behavior of the children exhibiting disruptive behavior was above 80% appropriate during baseline), (b) more than one teacher left the daycare, (c) more than one participant left the daycare, (d) retraining the teachers on the Level System occurred more than three times. None of these conditions was met. Specifically, the level of appropriate behavior exhibited by the children with disruptive behavior was 56.9% and 64.4%, no teachers or participants left the daycare, and retraining the teachers on the Level System did not occur.

Results

Effectiveness

The effectiveness of the strategies already used in the classroom and the Level System

was examined by visual and numerical inspection of the data obtained from behavioral observations (i.e., REDSOCS), the CGI, the Class Management Rating, and the Time-Out Log.

REDSOCS. The behavioral observation data were graphed after each observation and examined in four ways: (a) percentage of intervals of appropriate behavior for each target child per observation, (b) percent differences of appropriate behavior from one condition to another for each child, (c) average percentage of appropriate behavior across conditions for each child, and (d) overall averages for observation percentage, percent differences, and average percent across all children's data. Data were graphed averaged across teachers as well as individually. Percent differences were calculated by subtracting the percentage of behavior for one condition from the previous condition and then dividing by the percentage of behavior in the prior condition.

The percentages of appropriate behavior exhibited by individual participants in the classroom across teachers is displayed in Figure 1, and the percentage of the same behavior with individual teachers is displayed in Figures 2 through 5. The average percentages of appropriate behavior per condition per participant across teachers are depicted in Figure 6. The same data are presented by individual teacher in Figure 7. Average percentage data reported when considering child behavior with individual teachers should be accepted cautiously as many of these means are based on only two to four observations. When participants were absent from class, no data point appears on the Figures. Therefore, some data points were not connected.

Luke. Luke was considered to be a participant exhibiting disruptive behavior. He was observed to exhibit variable behavior throughout the conditions (see Figure 1). However, the percentage of his appropriate behavior was higher in the Level System conditions than in the baseline, withdrawal, and follow-up conditions. Additionally, during the withdrawal and follow-up phases, Luke's percentage of appropriate behavior was higher than it was during baseline.

These trends are made evident when the means of each condition are compared. Specifically, mean frequencies of appropriate behavior were 56.9% during baseline, 88.5% while using the Level System, 76.4% during the withdrawal phase, 91.6% during re-implementation of the Level System, and 74.9% during follow-up (Figure 6). These same general results were evident when Luke's appropriate behavior was inspected according to individual teachers (see Figure 2). However, mean levels of appropriate behavior were variable across different teachers (see Figure 7). Luke's appropriate behavior increased 55.4% during the Level System condition, decreased 13.7% during the withdrawal condition, increased 19.9% during re-implementation of the Level System, and decreased 18.1% during follow up (no Level System).

Cody. Cody also was considered to be a participant in the disruptive behavior category. Likewise, he was observed to exhibit variable behavior throughout the conditions (see Figure 1), and the percentage of his appropriate behavior was higher in the Level System conditions than in the other conditions. Additionally, during the withdrawal and follow-up phases, Cody's percentage of appropriate behavior was higher than it was during baseline. Mean frequencies of appropriate behavior for conditions A, B, A, B, and F were: 64.4%, 90.9%, 78.1%, 90.6%, and 82.5% respectively (Figure 6). These same general results were evident in Cody's appropriate behavior when he was in Colleen's and Patricia's classes (see Figure 3). However, in Rachel's class, the percentage of appropriate behavior that he exhibited during baseline was comparable to the appropriate behavior during withdrawal (Figure 3). When considering his appropriate behavior in Chad's class, there was significant overlap in the amount of appropriate behavior exhibited in the Level System conditions and the withdrawal condition (Figure 3). In other words, when the Level System was withdrawn, his behavior did not change. In Lori's class, Cody engaged in more appropriate behavior during follow up, with no Level System in place,

than in any other condition (Figure 3). These differences are evident when mean levels of appropriate behavior across teachers are compared (see Figure 7). Cody's appropriate behavior increased 41.3% during the Level System condition, decreased 13.7% during the withdrawal condition, increased 16.1% during re-implementation of the Level System, and decreased 9 % during follow up (no Level System).

Tim. Tim was a participant exhibiting typical behavior. The percentage of appropriate behavior that he exhibited was highest in the Level System and withdrawal conditions (including follow up) (Figures 1 and 6). Additionally, his behavior became more stable after baseline. Mean frequencies of appropriate behavior for each condition were as follows: 94.9% (baseline), 99.2% (Level System), 99.1% (withdrawal), 99.4% (Level System), 98.5% (follow up). Similar results were found when considering each teacher's class individually (see Figures 4). The similar patterns of results across teachers are evident when mean levels of appropriate behavior across teachers are compared (see Figure 7). Tim's percentage of appropriate behavior increased 4.5% during implementation of the Level System, decreased 0.04% when the Level System was withdrawn, increased 0.29% when the Level System was re-introduced to the classes, and decreased 0.97 % at 1-month follow up (no Level System).

John. John also was a participant considered to be exhibiting typical behavior. He exhibited more appropriate behavior in the Level System conditions and in the withdrawal and follow-up conditions than during baseline (Figures 1 and 6). Furthermore, his behavior became more stable during the Level System conditions when compared to the other conditions. Figure 6 depicts trends with a comparison of the means of each condition. Specifically, mean frequencies of appropriate behavior were 95.5% during baseline, 100% while using the Level System, 98.6% during the withdrawal phase, 99.2% during re-implementation of the Level System, and 99.2%

during follow up. Similar results were found when considering each teacher's class individually (see Figure 5). Although, when considering John's behavior while in Lori's class (Figure 5), the percentage of appropriate behavior exhibited during the withdrawal condition was below that which it was during all other conditions, including baseline. Figure 7 displays mean levels of appropriate behavior across teachers. John's percentage of appropriate behavior increased 4.7% during the Level System condition, decreased 1.4% during the withdrawal condition, increased 0.6% during the second Level System condition, and decreased 0.07 % during the 1-month follow up condition (no Level System).

CGI. Raw scores on the CGI were converted to age- and gender-based T-scores for interpretation. T-scores were obtained at the beginning of the study and after each condition and examined in a similar manner to the behavioral observation data (e.g., averages, percent differences). Data were obtained across teachers and target children as well as individually. Additionally, CGI data were averaged across all children in the class and examined.

Mean T-scores for each target child for each condition are displayed in Table 1. Additionally, the same data were considered for each teacher individually (see Table 1). As discussed previously, T-scores between 45 and 55 were considered typical and T-scores above 60 were considered to be in the disruptive range.

Luke. Overall, Luke's T-scores averaged across teachers indicated that even though these scores decreased during the first Level System condition, they always remained in the disruptive range (70.4-59.8) (Table 1). This trend also is evident when considering Patricia and Colleen's individual data. However, during the first B condition, Luke's T-scores were in the typical range according to these two teachers (T-score = 53). No other systematic changes were evident across conditions. Luke's mean T-scores decreased 8.5% during the baseline condition, decreased 7.1%

during the Level System condition, increased 8.3% during the withdrawal condition, decreased 0.62% during re-implementation of the Level System, and increased 0.62% during follow up (no Level System).

Cody. Cody's T-scores averaged across teachers were relatively stable throughout the study. However, his mean T-score decreased during the follow-up condition to within typical limits (T-score = 53.6) (Table 1). Cody's T-scores by individual teachers were more variable, except for Chad's ratings which were similar to overall ratings. For example, according to Patricia, Cody's T-score decreased during baseline to within typical limits and then increased steadily until follow up (T-score = 44). Colleen rated Cody's behavior as in the typical range throughout all conditions except withdrawal. According to Lori's ratings, Cody's T-scores were in the disruptive range during all conditions. Also, his T-scores were the highest during selection and the withdrawal phase. Cody's T-scores from Rachel indicate that they increased during baseline and then decreased throughout the rest of the study. Cody's mean T-scores on the CGI decreased 5.6% during baseline, increased 2.5% during the Level System condition, increased 5.1% during the withdrawal condition, decreased 2.9% during re-implementation of the Level System, and decreased 10.1 % during follow up (no Level System).

Tim. Tim's mean T-scores across teachers were relatively stable throughout all conditions (Table 1). Scores always were in the typical range, and no systematic changes were evident. Similar results were found when considering each teacher's class individually. Tim's mean T-scores increased 0.44% during baseline, decreased 1.8% during implementation of the Level System, decreased 1.8% when the Level System was withdrawn, decreased 3.2% when the Level System was re-introduced to the classes, and increased 1.7 % at 1-month follow up (no Level System).

John. Similar to Tim, John's mean T-scores indicate no systematic changes across conditions, and scores remained in the typical range (Table 1). Similar results are found when considering each teacher's scores individually. John's mean T-scores decreased 6.2% during baseline, decreased 0.47% during the Level System condition, remained unchanged during the withdrawal condition, increased 0.95% during the second Level System condition, and remained unchanged during the 1-month follow-up condition (no Level System).

All children in the class. No systematic changes across conditions were evident when considering all of the children in the class across teachers or by individual teachers. Mean T-scores ranged from 50.4 to 48.4.

Class Management Rating. The Class Management Rating was completed after each class. The rating was a 5-point Likert-type scale in which "1" indicated extreme unmanageability and "5" indicated extreme manageability. Table 2 indicates class management ratings of teachers combined, and by individual teachers. Mean class management ratings suggest that, although relatively stable throughout the study, ratings of class management increased steadily and then decreased slightly at follow up. Specifically, mean ratings were as follows: 4.3 during baseline, 4.4 during implementation of the Level System, 4.5 during withdrawal, 4.8 during re-introduction of the Level System, and 4.6 during follow up. Overall, class management ratings across teachers increased 4.1% in the Level System condition, increased 1.2% in the withdrawal condition, increased 6.1% in the second Level System condition, and decreased 4.5% in the follow-up condition.

Colleen's ratings follow the same general trend (see Table 2). Patricia's ratings of the manageability of her class decreased during the first Level System condition (4.8 to 4.7) and then increased to ceiling levels for the entirety of the study (5 for each condition). Chad's class

management ratings were highest during the first Level System condition (4.3), and then decreased (3.9) and remained relatively stable throughout the study (3.8 and 3.7 for each condition). Lori's class management ratings decreased slightly when the Level system was removed from her class (4.3 to 4.1), and increased to ceiling levels with re-implementation (5). Rachel's class management ratings increased during the withdrawal phase (4.3 to 4.7) and continued to increase until the end of the study (4.9 and 5, respectively).

Time-Out Log. Similar to the other effectiveness measures, data from the time-out log included percent differences among conditions, average score across conditions, and overall averages. Data were analyzed with all teachers combined as well as by individual teacher. Additionally, data were obtained and examined concerning target children as well as all children in the class. Time-out logs were completed by the teacher after each class.

The mean frequency of time outs given to target participants with disruptive behavior across teachers and by individual teachers is displayed in Table 3. Because no time outs were given to the participants with typical behavior throughout the duration of the study, these data were not reported. Additionally, Table 3 includes the mean frequency of time outs given to all children in the class (including target participants) across teachers and by individual teachers.

Luke. Luke was given between 0 and 4 time outs per observation day. However, the number of days in which he received time outs decreased as the study progressed. When considering the means for each condition (see Table 3), it is clear that the number of time outs that he received were fewer during the Level System conditions and follow up. Specifically, mean number of time outs given to Luke per condition were as follows: 0.76 (condition A), 0.39 (condition B), 0.54 (condition A), 0.2 (condition B), 0.06 (follow up). The number of time outs for Luke decreased throughout the study when considering individual teachers. The mean

number of time outs given by Patricia, Lori, and Rachel provide evidence for this trend. The number of time outs given by Colleen and Chad increased before they decreased. The overall frequency of time outs given to Luke decreased 48.7% during the Level System, increased 38.5% during withdrawal, decreased 63% during re-implementation of the Level System, and decreased 70% during follow up (no Level System).

Cody. Similarly, Cody received the fewest time outs during the Level System conditions and follow up (Table 3). Cody was provided between 0 and 3 time outs per observation day. The mean number of time outs given to Cody was 0.36 during baseline, 0.17 during the Level System condition, 0.46 during withdrawal, 0.1 during re-introduction of the Level System, and 0.12 during the follow up condition (no Level System). These same trends are apparent when examining the data by individual teacher. The number of time outs that Cody received decreased 52.8% when the Level System was in place, increased 170.6% when the Level System was removed, decreased 78.3% when the Level System was put back in place, and increased 20% during follow up, when no Level System was in place.

All children in the class. Table 3 reveals the mean frequency of time outs that were given to all of the children in the class during each condition across all teachers and by individual teacher. Data indicate that more time outs were given during the conditions when the Level System was not being used, except during follow up. Specifically, the average number of time outs was as follows for each condition: 2 (baseline), 1.1 (Level System), 1.5 (withdrawal), 0.35 (Level System), 0.24 (follow up). These same trends are apparent when considering the mean number of time outs provided by individual teachers. However, the data suggest that the mean number of time outs given by Lori does not follow this trend. Instead, the number of time outs that she gave steadily decreased throughout the study. Overall, the number of time outs provided

to the children in the class decreased 48.3% during Level System, increased 42.1% during withdrawal, decreased 76.7% during Level System, and decreased 32.8% during follow up.

Satisfaction

Teacher treatment satisfaction was determined by calculating and interpreting the scores on the IRP at the end of each condition, and comparing the scores between conditions.

Whichever condition resulted in higher scores on the measure, and whichever management strategies that the teachers used after the last experimental session was considered the system in which the teachers were the most satisfied.

Parent treatment satisfaction was determined by a parental interview that assessed the parent's acceptability of the Level System and the behavioral principles that are used with the Level System. High scores on the interview would indicate satisfaction with the Level System.

Intervention Rating Profile (IRP). The highest possible score on the IRP is a 120. Higher scores indicate greater acceptability of the intervention. The mean satisfaction rating increased slightly between baseline and the first Level System condition, and the second Level System condition and 1-month follow up. A slight decrease in satisfaction was apparent between the first Level System condition and withdrawal, and the withdrawal condition and the second Level System condition. Specific average scores on the IRP for each condition were as follows: 88.8 (baseline), 97.2 (Level System), 87.8 (withdrawal), 86.4 (Level System), and 88.4 (follow up). When considering satisfaction ratings of individual teachers, Colleen's ratings were similar to overall ratings (92, baseline; 108, Level System; 106, withdrawal; 103, Level System; 109, follow up). However, Chad's satisfaction rating increased during implementation of the Level System (85 to 98), but decreased for the remainder of the study (90, 78, 73), and Lori's ratings decreased steadily throughout the study (92, baseline; 89, Level System; 70, withdrawal; 61,

Level System; 64, follow up). Additionally, Patricia's scores indicate that she was more satisfied with the Level System (109 and 112) than her regular class management strategies (92 and 106), and Rachel's scores remained relatively stable throughout each condition (83, baseline; 82, Level System; 80, withdrawal; 78, Level System; 81, follow up). Despite overall and individual teacher (Colleen, Chad, and Patricia) reported level of satisfaction with the Level System during the first implementation, none of the teachers chose to use the Level System at follow up rather than their typical classroom management strategies.

Parent interview. The parental interview was completed at the end of the study by willing parents of child participants in the study. All of the parents completed the measure (13 parents). Five of the parents completed the interview over the telephone with the experimenter, and eight of the parents completed the interview in a face-to-face format at the daycare. Higher scores indicate more acceptability of behavioral management strategies used with the Level System (e.g., positive reinforcement, token economy, response cost). In addition to scored items, this measure included five open-ended items in which parents were asked to list positive and negative concerns regarding the Level System and time out.

Results from the parent interview indicate that the parents were very accepting of use of the Level System in their child's class. Specifically, the mean total score was 52.8 out of 60 ($SD = 5.7$). On the open-ended items, parents described positive aspects of the Level System to include: rewards, visual learning, and motivation to behave appropriately. Parental concerns regarding the Level System included: competition among children, embarrassing to parents because the Level System is posted for others to see, and embarrassing to children if they do not receive the reward. Concerning time out, parents reported positive aspects to include: modeling of consequences, and provides children time to contemplate their inappropriate behavior. Parents

described concerns with time out to include the following: may last too long, and may positively reinforce inappropriate behavior.

Negative Effects

To determine if the Level System resulted in any negative effects on “intrinsic motivation,” the appropriate behavior data of the participants were graphed and visually inspected to determine if during the withdrawal or withdrawal condition, the children’s rate of appropriate behavior decreased below baseline levels, then increased to baseline levels and remained stable. Additionally, mean percentages of appropriate behavior were compared to determine if the percentage during withdrawal was less than the percentage during baseline. If the rate of appropriate behavior did not decrease below baseline levels, then it would be determined that no negative effects on “intrinsic motivation” were evident from the Level System.

Visual inspection of graphed data as well as comparing mean percentages of appropriate behavior during baseline and withdrawal suggest that no negative effects on “intrinsic motivation” were evident. Specifically, when considering each participant’s overall and mean level of appropriate behavior across teachers (Figures 1 and 6) and in individual teacher’s classes (Figures 2 to 5, and 7), no negative effects were revealed, except for John when he was in Lori’s class (see Figures 5 and 7). John’s level of appropriate behavior during the withdrawal condition was below the level during baseline. Specifically, John’s mean percentage of appropriate behavior was 95.8 during baseline and 93.9 during withdrawal. In other words, the percentage of appropriate behavior that John exhibited decreased 2% from baseline to withdrawal. However, no other negative effects on “intrinsic motivation” were found.

Treatment Integrity

The data from the behavioral observations of teacher response to child behavior generally support the results from the treatment integrity checklist, indicating that the teachers were using the Level System with integrity. Table 4 indicates that when the participants were behaving appropriately, appropriate teacher responses were provided (e.g., praise, moving shapes up, no time outs). Also, these behaviors increased when the Level System was being used compared to when it was not being implemented. Furthermore, when the participants engaged in inappropriate behavior, appropriate teacher responses were provided (e.g., warnings, moving shapes down, time outs) (see Table 4). However, some results from the teacher behavioral observations suggest that the teachers were not using the Level System with excellent integrity. For example, teachers occasionally provided participants with praise when their behavior was coded as inappropriate. In addition, although the teachers were providing the children with adequate responses to behavior (e.g., warnings for inappropriate behavior), the percentage of time that they engaged in these responses could be considered low (e.g., 26.2% warnings for inappropriate behavior).

Discussion

The current study was conducted to examine the effectiveness, satisfaction, and potential negative effects of the Level System compared to typical classroom management strategies in managing disruptive behavior in preschool classrooms. Specifically, the current study evaluated whether or not the children's appropriate behavior increased while the Level System was used in the classroom, as compared to the strategies already utilized by the teachers. In addition, satisfaction with the intervention was assessed by teacher and parent report of satisfaction as well as which approach the teachers chose to continue after termination of the study. Lastly, possible negative effects on "intrinsic motivation" were examined by visual inspection of graphic

representations of the data, and mean comparisons of data across conditions.

In general, there were four main findings in the current study. First, when examining data from children with disruptive behavior and children with typical behavior, it appears that the Level System is more effective in managing disruptive behavior than previously-used classroom management strategies. Second, the data suggest that fewer time outs were given while the Level System was used in the class compared to while the typical classroom management strategies were used. Third, teacher report of satisfaction with the Level System varied. However, the parents reported high levels of satisfaction with the Level System. Lastly, negative effects on “intrinsic motivation” with use of the Level System were not evident. Several methodological limitations should be considered when interpreting these results. Discussion concerning the main findings, methodological limitations, clinical implications, and directions for future research follows.

Main Findings

Overall Effectiveness of the Level System

It was expected that the Level System would result in more appropriate behavior exhibited by children with disruptive behavior and by children with typical behavior in the classroom when compared to conditions in which the Level System was not used. Visual inspection of behavioral observation (i.e., REDSOCS) data provides support for this expectation. When the participants were considered separately and together across teachers and in individual teacher’s classes, percentages of appropriate behavior were higher in the Level System conditions than in the other conditions. However, these results are less clear when considering the children with typical behavior because ceiling effects made it difficult to distinguish differences among conditions. Data also appeared to be more stable during the Level System

conditions when compared to the conditions that utilized typical classroom management strategies. This may have occurred because the teachers used a more consistent behavioral management program compared to their typical strategies. For example, the Level System requires that the teachers provide consequences each time that a child engages in inappropriate behavior, whereas there is no such requirement in a typical classroom.

On teacher report measures, the effectiveness of the Level System was less evident. For example, visual inspection and mean comparison of class management ratings suggest that the teachers perceived themselves to be better able to manage their class as the study progressed, except at the 1-month follow up. This result may have occurred because the teachers felt better able to manage their class as the school year continued. Another possible explanation is that the teachers felt more confident managing their classes when they received feedback concerning their management behavior. Then, during follow up, after 1 month without feedback from the experimenter, they felt less confident managing their classes. Furthermore, on the CGI, significant treatment effects were not evident. Specifically, no changes in CGI scores were apparent for the children with disruptive behavior or the children with typical behavior.

The lack of convergence between teacher-report measures and behavioral observation data may have occurred because the teachers were not sensitive to the behavior change that was evident by behavioral observation (Bahl et al., 2000). A potential explanation for these findings is that the CGI measures behaviors that the Level System did not address (e.g., fails to complete tasks, short attention span). Alternatively, it is possible that the teachers “labeled” children as challenging early in the school year and had difficulty noticing the positive behavior improvements because of the negative effects of the label (Polyson, 1979). Finally, it is important to consider the possibility that improvements evident on a highly sensitive behavioral

observation measure were not as noticeable in the children's daily behavioral functioning.

Overall, these results suggest that the Level System resulted in more appropriate behavior exhibited by children with typical and disruptive behavior in the class. However, the changes in behavior may not be as evident to the teachers. Thus, only preliminary support is provided to the effectiveness of the Level System. In order to ensure that teachers will use strategies that are found to be effective in the classroom, it is pertinent that measures of change be used and provided to teachers that evaluate behaviors that are important to teachers so that changes in these behaviors are more evident to teachers.

Time Out

Generally, the number of time outs given by the teachers was lower when the Level system was used than when typical classroom strategies were used, except for follow up. The number of time outs provided to the children was lowest during follow up. This trend was evident among the children with disruptive behavior as well as all of the children in the class. For the most part, this trend was clear when considering the number of time outs given by individual teachers. The frequency of time outs given to the children probably decreased during the Level System conditions because the teachers were taught to use the system to manage minor misbehaviors and to prevent minor misbehaviors from escalating to more severe disruptive behavior. However, this does not explain the finding that the fewest time outs were provided during follow up. Potentially, this result might be due to generalization of the behavioral management skills that were taught with the use of the Level System (e.g., positive reinforcement, differential attention, warnings) after the Level System was removed.

Satisfaction

Parents appeared to be satisfied with the Level System. Specifically, no parents reported

that their child would be humiliated or embarrassed by having their name in the “cloudy zone” of the token economy, only by not receiving the reward. Additionally, most parents stated that they would not be embarrassed if their child’s name was in the “cloudy zone.” This was, however, a concern for some parents. Overall, parents described the Level System positively.

Teacher satisfaction with the Level System was not as evident as parent satisfaction. Specifically, the teachers appeared to be more satisfied with the Level System than they were in the baseline phase; however, after removal of the initial implementation of the Level System, teachers reported more satisfaction when the Level System was not in place. This is not surprising since the percentage of appropriate behavior exhibited by the participants did not decrease to baseline levels during the withdrawal phase. In other words, the participants behaved more appropriately during the withdrawal phase than baseline, therefore, the teachers may not have considered re-implementation of the Level System necessary. On the other hand, if the children’s behavior had returned to baseline levels, the teachers may have considered re-implementation of the Level System to be worth the time and effort. This reason also may account for the teachers choosing not to use the Level System during the follow-up assessment. Another potential explanation for this finding may be that once the teachers learned more effective behavior management strategies (e.g., positive reinforcement) from using the Level System, benefits gained from adding the Level System were not worth the time and effort of implementation. The teachers may have learned skills (e.g., praise) that allowed them to better manage their classrooms without the use of a token economy. Teachers reported to the experimenter that they used the skills that were taught to them, but that they felt that using the Level System took too much time and was too difficult without a stable teacher’s aide. It should be noted that for the current study, teachers were required to complete several daily

questionnaires and forms (e.g., class management rating, time out log) that would not be required when using the Level System clinically. Teachers reported that they would have been more likely to use the Level System if additional forms were not required.

Negative Effects

Because the children's level of appropriate behavior did not decrease below baseline levels during the withdrawal phase, it was determined that no negative effects on "intrinsic motivation" were evident with implementation and use of the Level System. Individually, only one child in one class potentially met the current study's criteria for loss of "intrinsic motivation" and that was in the presence of only 1 of the 5 teachers. Because no other children (considered individually or combined) met this study's criteria in the presence of any teacher (considered individually or combined), the current study supports previous research that token economies do not result in a loss of "intrinsic motivation" (e.g., Cameron et al., 2001). Further support is added to this result because of the length of the withdrawal phase. Specifically, the withdrawal phase was held for stability (i.e., 15 observations for each child) plus 1 week. Previous studies may have found a "loss of intrinsic motivation" because the length of the withdrawal phase was very short. Behaviorists often predict a decrease in behavior below baseline immediately following the removal of a reward program because of satiation with the reinforced behavior or a negative contrast effect (e.g., Cameron et al.; Reitman, 1998). The longer withdrawal phase was included in this study to allow time for temporary effects to stabilize. In the current study, however, the percentage of appropriate behavior exhibited by the participants did not even return to baseline levels, such that the longer withdrawal was not necessary.

Methodological Limitations

Participant Selection

It appeared that teacher report on the CGI for selection and behavioral observations coincided when considering the children with disruptive behavior. However, inconsistencies were found between teacher report for selection and behavioral observations for the children with typical behavior. This is evident because the children that the teacher's reported as engaging in typical levels of appropriate behavior were observed to exhibit extremely high rates of appropriate behavior during baseline. As a result, a ceiling effect occurred and made it difficult to distinguish differences among conditions when considering the appropriate behavior of the children with typical behavior.

Observation Assessment

REDSOCS. The observation assessment that was used for the current study produced approximately 15 min of data on each child per observation. Thus, only a limited amount of data was collected for each participant which may have contributed to the variability in the children's behavioral data. It is important to note, however, that clear differences were evident across conditions even with the use of limited observations.

Teacher response data. One measure of treatment integrity used in the current study was behavioral observation of teacher response to child behavior. This measure may have underestimated the amount of integrity with which the teachers used the Level System. Teacher responses were coded toward the particular target child being observed at the time. Therefore, any teacher response to another child's behavior was not coded. For example, a teacher may have been moving several children's shapes up for appropriate behavior, but if these children were not being observed, this adherence would not have been recorded. Therefore, the

percentage of time that the teachers engaged in the integrity behaviors may have been underestimated.

Because the data collected concerning teacher behavior are preliminary and the codes were created for the current study, the results are difficult to interpret. The percentage of time that the teachers should be spending engaging in each of the integrity behaviors to consider the Level System as implemented accurately is unclear. For example, even if the teachers were implementing the Level System perfectly, it would be unreasonable for them to praise and move shapes up 100% of the time for appropriate child behavior, and warn and move shapes down 100% of the time for inappropriate child behavior. If this were to occur, the teachers would not meet their teaching requirements or attend to the behavior of other (non-target) children in the class. Thus, teacher response data (e.g., greater use of praise for appropriate behavior during Level System conditions) provide some support that the Level System was used with integrity throughout the study, but to what extent is uncertain.

Experimental Design

A significant limitation of the current study is that stability was not reached for all conditions for either of the children with disruptive behavior when considering their behavior across teachers. Stability, however, was reached for all conditions for the children with typical behavior, but, as stated previously, this is a limitation itself because ceiling effects occurred. When inspecting the children with disruptive behavior's data with individual teachers, stability was reached on a few occasions (e.g., Cody's baseline data with Patricia), however, no clear trends are present. It may have been difficult to achieve stability with the current population. Specifically, preschool child behavior is variable and depends on numerous factors (e.g., sleep, breakfast). Additionally, the environment of the current study may have made it difficult to

achieve stability because the children rotated through 2 teachers each day, and their behavior may have been different in each teacher's class. To circumvent lack of stability from affecting the results of the current study as much as possible, each condition was held for a minimum of 17 days.

Another limitation concerning the experimental design of this study is that the participants' behavior did not reverse to baseline levels during the withdrawal condition or follow up assessment (i.e., no Level System). This carryover effect, or incomplete withdrawal (Parsonson & Baer, 1986) obfuscates the effects of the intervention on the children's behavior. A carryover effect is "characterized by the experimenter's inability to retrieve original levels of baseline responding" (Barlow & Hersen, 1984, p. 99). Researchers suggest (e.g., Barlow & Hersen) that carryover effects occur for several reasons. For example, they may occur due to changes in instructions concerning the intervention across conditions (Kazdin, 1973), new conditioned reinforcers that are established (Bijou et al., 1969), differences in stimuli across conditions (Kazdin & Bootzin, 1972), and naturally occurring contingencies in the environment that may maintain the new behaviors (Krasner, 1971). To circumvent carryover effects, Bijou et al. recommend shortening the intervention conditions so that new conditioned reinforcers are less likely to be established. However, as stated previously, in the current study, conditions were held longer in an attempt to obtain stability. A possible explanation for the carryover effect in the current study is that the teachers learned to use skills with the Level System that could have been used when the Level System was removed. For example, the teachers were taught to use labeled praise when implementing the Level System, and labeled praise can be used without the Level System in place. Therefore, the teachers' behavior (i.e., labeled praise) "carried over" affecting child levels of appropriate behavior>

Teachers' Aides

The class of children in the study had a rotating teacher's aide. Specifically, there was an aide present on random occasions. Additionally, the aide changed often. At the beginning of the intervention condition, the current teachers' aide on staff was trained in the use of the Level System, and data were to be collected concerning the aide's implementation of the Level System. However, once it was discovered that the teachers' aide was not always present (i.e., she attended the class approximately 1 day per week), and shortly after the implementation of the Level System (5 observations), the teacher's aide changed (i.e., a new aide was hired), it was determined that this data would be sporadic. Therefore, two future teachers' aides were not trained in the use of the Level System, and data were not collected concerning their behavior.

During the first 5 observations of the first Level System condition, the teachers' aide was using the Level System. However, because she did not attend the class every day, she was in the class on the first observation of this condition. From visual inspection, it appears that her presence with using the Level System did not affect the participants' behavior.

Setting

The environment in which the study was conducted is another limitation of the current study. Specifically, the setting was a privately-owned preschool that served mostly middle-class families. Thus, similar results may not have been found if the study was conducted in a Head Start classroom in which the children are at a higher risk for disruptive behavior problems. Additionally, different results may have been found with teachers that were not as well-trained as those that participated in this study. All of the teachers in the current study received bachelor's degrees, and had several years of teaching experience. Because of the setting of the preschool, the teachers interacted with each other on a daily basis. Therefore, it is likely that discussions

took place concerning the Level System. Some teachers may have decided not to use the Level System at the study's end because other teachers were not using the system.

Clinical Implications

Several clinical implications are apparent from the current research study. First, it is interesting that when using the Level System the number of time outs given by the teachers decreased. This result may be due to an increased number of low-level consequences (i.e., moving children's shapes down a level) for minor misbehavior. These consequences may have prevented the children's behavior from escalating to disruptive behavior requiring a higher level consequence of time out. This is an important consideration when determining which behavior management strategies to use in the classroom. Specifically, some parents or teachers may not approve of using time out procedures because they may require more physical guidance. Therefore, using a whole-class token economy such as the Level System may circumvent using higher level consequences for disruptive behavior.

Another clinical implication is evident when considering the teachers' satisfaction with the Level System. The teachers appeared to be more satisfied with the Level System when they first began using it than when they used it in their classroom a second time. This finding implies that teachers may be satisfied using the Level System for a short period to learn to better manage their classroom, but not using the token economy for an extended length of time. Implementation of the Level System resulted in a 47.9% increase in appropriate behavior from the children exhibiting disruptive behavior. This increase appears to be clinically significant as evidenced by the teachers' increase in satisfaction with the first implementation. Therefore, it may be better to use the Level System in the classroom for disruptive periods, and then fade out use of the system. For example, teachers could implement the Level System at the beginning of the school

year to teach children class rules and expectations within the classroom, and then fade the system out once natural contingencies (e.g., positive reinforcement from peers, access to activities) take over. Additionally, the teachers reported that they were not completely satisfied with the Level System because they felt that it took time away from teaching. Therefore, using the Level System in the classroom may be more appropriate when a steady teacher's aide is available to assist with behavior management.

Lastly, this study provides evidence that the Level System resulted in no negative effects on preschool children's "intrinsic motivation." Additionally, the Level System was proven to be effective in increasing appropriate behavior exhibited by both children with disruptive behavior and children with typical behavior. Therefore, introducing a whole-class token economy, such as the Level System, should be considered by clinicians and school consultants at the preschool level when referrals are made concerning children with disruptive behavior problems.

Directions for Future Research

This study provides preliminary support for using a whole-class token economy, the Level System, to manage disruptive behavior in a preschool classroom. Thus, future research should confirm and expand upon the current findings. Specifically, because the current evaluation of the Level System was conducted with one class across five teachers, it is important to evaluate this system with numerous classrooms each with a different teacher. This would provide evidence as to the generalizability of the current results. Furthermore, the Level System should be evaluated with diverse populations. For example, determining the effectiveness of this token economy with children with developmental disabilities, learning disabilities, and other mental health concerns (e.g., ADHD) would aid in determining the applicability of the Level System with diverse preschool children.

A series of systematic examinations of the specific components of the Level System may yield valuable information. These investigations could examine if both the token economy and the response cost components of the Level System are required to obtain significant changes in appropriate behavior exhibited by participants. Additionally, it would be interesting to evaluate the effect of the addition of a structured time out procedure to the Level System to conclude if similar results would be obtained with regard to the number of time outs given by the teachers compared to the results of the current study. Another component that should be evaluated is the visual display of the chart. Specifically, given that some parents expressed that they would be embarrassed if their child's name was in the cloudy zone, research should address the issue of the need for the chart to be displayed during use of the Level System to obtain changes in levels of appropriate behavior.

Other future research should focus on determining if a whole-class token economy, such as the Level System, would be beneficial and produce similar results during playground activities. Throughout the current study, teachers requested the need for such a system to use during outside time. Additionally, because teacher satisfaction with the Level System was inconsistent, it should be researched further. If teachers are not completely satisfied with the time and effort involved in using the Level System, then, perhaps it only should be implemented when a teacher's aide is available for implementation. Alternatively, it may be possible to make modifications to the system to make it less time consuming (e.g., have children move their pieces up and down, use it only during transitions, provide rewards less frequently). Thus, future research is needed before conclusions can be drawn about the general efficacy, practicality and ecological validity of a whole-classroom token economy for managing preschool behavior problems.

References

- Anhalt, K., McNeil, C. B., & Bahl, A. B. (1998). The ADHD Classroom Kit: A whole-classroom approach for managing disruptive behavior. *Psychology in the Schools, 35*, 67-79.
- Bacharach, V. R., & Baumeister, A. A. (2003). Child care and severe externalizing behavior in kindergarten children. *Journal of Applied Developmental Psychology, 23*, 527-537.
- Bahl, A. B., McNeil, C. B., Cleavenger, C. J., Blanc, H. M., Bennett, G. M. (2000). Evaluation of a whole-classroom approach for the management of disruptive behavior. *Proven Practice, 2*, 62-71.
- Baker, J. G., Stanish, B., & Fraser, B. (1972). Comparative effects of a token economy in nursery school. *Mental Retardation, 10*(4), 16-19.
- Barkley, R. A. (1987). Poor self-control in preschool hyperactive children. *Medical Aspects of Human Sexuality, 21*, 176-180.
- Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change* (2nd ed.). New York: Pengamon Press Inc.
- Berger, K. S. (1998). *The developing person through the lifespan* (4th ed.). New York: Worth Publishers.
- Bijou, S. W. et al. (1969). Methodology for experimental studies of young children in natural setting. *Psychological Record, 19*, 177-210.
- Brown, R. T., & Wynne, M. E. (1982). Correlates of teacher ratings, sustained attention, and impulsivity in hyperactive and normal boys. *Journal of Clinical Child Psychology, 11*, 262-267.

- Budd, K. S., Leibowitz, J. M., Riner, L. S., Mindell, C., & Goldfarb, A. L. (1981). Home-based treatment of severe disruptive behaviors: A reinforcement package for preschool and kindergarten children. *Behavior Modification, 5*, 273-298.
- Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior Analyst, 24*, 1-44.
- Campbell, S. G. (1990). *Behavior problems in preschool children: Clinical and developmental issues*. New York, NY: The Guilford Press.
- Clark, M. M., & Cheyne, W. M. (Eds.). (1979). *Studies in preschool education*. London: Hodder and Stoughton.
- Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126*, 884-888
- Conners, C. K. (1997). *Conners' Rating Scales - Revised*. New York: Multi-Health Systems Inc.
- Cooper, J. O. Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Corrigan, P. W. (1995). Use of token economy with seriously mentally ill patients: Criticisms and misconceptions. *Psychiatric Services, 46*, 1258-1263.
- Davidson, P., & Bucher, B. (1978). Intrinsic interest and extrinsic reward: The effects of a continuing token program on continuing nonconstrained preference. *Behavior Therapy, 9*, 222-234.
- Diamond, J. M., & Deanne, F. P. (1990). Conners' Teachers' Questionnaire: Effects and implications of frequent administration. *Journal of Clinical Child Psychology, 19*, 202-204.

- Dickinson, A. M. (1989). The detrimental effects of extrinsic reinforcement on “intrinsic motivation.” *The Behavior Analyst, 12*, 1-15.
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Boston, MA: Allyn and Bacon.
- Drege, P., & Beare, P. L. (1991). The effect of a token reinforcement system with a time-out backup consequence on the classroom behavior of E/BD students. *British Canadian Journal of Special Education, 15*, 39-46.
- Dunifon, R., Kalil, A., & Danziger, S. K. (2003). Maternal work behavior under welfare reform: How does the transition from welfare to work affect child development? *Children and Youth Service Review, 25*(1-2), 55-82.
- DuPaul, G. J. (1991). Parent and teacher ratings of ADHD symptoms: Psychometric properties in a community based sample. *Journal of Clinical Child Psychology, 20*, 245-253.
- DuPaul, G. J., & Barkley, R. A. (1992). Situational variability of attention problems: Psychometric properties of the Revised Home and School Situations Questionnaire. *Journal of Clinical Child Psychology, 21*, 178-188.
- Elliott, S. N., Witt, J. C., Galvin, G. A., & Peterson, R. (1984). Acceptability of positive and reductive behavioral interventions: Factors that influence teachers' decisions. *Journal of School Psychology, 22*, 353-360.
- Epstein, L. H., Masek, B. J., & Marshall, W. R. (1978). A nutritionally based school program for control of eating in obese children. *Behavior Therapy, 9*, 766-778.
- Epstein, M. H., & Nieminen, G. S. (1983). Reliability of the Conners' Abbreviated Teacher Rating Scale across raters and across time: Use with learning disabled students. *School Psychology Review, 12*, 337-339.

- Filcheck, H. A., McNeil, C. B., Greco, L. A., & Bernard, R. S. (in press). Using a whole-class token economy in a preschool classroom to manage disruptive behavior. *Psychology in the Schools*.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: John Wiley and Sons.
- Ford, J. D., & Foster, S. L. (1976). Extrinsic incentives and token-based programs: A reevaluation. *American Psychologist, 31*, 87-90.
- Hembree-Kigin, T. L., & McNeil, C. B. (1995). *Parent-Child Interaction Therapy*. New York: Plenum Press.
- Herman, S. H., & Tramontana, J. (1971). Instructions and group versus individual reinforcement in modifying disruptive group behavior. *Journal of Applied Behavior Analysis, 4*(2), 113-119.
- Hoffman, J. B., & DuPaul, G. J. (2000). Psychoeducational interventions for children and adolescents with attention-deficit/hyperactivity disorder. *Child and Adolescent Psychiatric Clinics of North America, 9*, 647-661.
- Jacobs, J. R., Boggs, S. R., Eyberg, S. M., Edwards, P. D., Querido, J. G., McNeil, C. B., & Funderburk, B. W. (2000). Psychometric properties and reference point data for the revised edition of the School Observation Coding System. *Behavior Therapy, 31*, 695-712.
- Jones, M. L., Eyberg, S. M., Adams, C. D., Boggs, S. R. (1998). Treatment acceptability of behavioral interventions for children: An assessment by mothers of children with disruptive behavior disorders. *Child and Family Behavior Therapy, 20*(4), 15-26.

- Jones, R. N., Downing, R. H., Latkowski, M. E., Ferree, R. C., & McMahon, W. M. (1992). Level systems as shaping and fading procedures: Use in a child inpatient psychiatry setting. *Child and Family Behavior Therapy, 14*(2), 15-37.
- Kazdin, A. E. (1973). Methodological and assessment considerations in evaluating reinforcement programs in applied settings. *Journal of Applied Behavior Analysis, 6*(3), 517-531.
- Kazdin, A. E. & Bootzin, R. R. (1972). The token economy: An evaluative review. *Journal of Applied Behavior Analysis, 5*(3), 343-372.
- Kelley, M. L., Heffer, R. W., Gresham, F. M., & Elliot, S. N. (1989). Development of a modified treatment evaluation inventory. *Journal of Psychopathology and Behavioral Assessment, 11*, 235-247.
- Kohn, A. (2000). *The schools our children deserve: Moving beyond traditional classrooms and "tougher standards."* Boston, MA: Houghton Mifflin Company.
- Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes.* Boston, MA: Houghton Mifflin Company.
- Krasner, L. (1971). Behavior therapy. *Annual Review of Psychology, 22*, 483-532.
- Kysela, G. M. (1972-1973). A token economy. *Mental Retardation Bulletin, 1*(2), 33-36.
- Lavigne, J. V., Gibbons, R. D., Christoffel, K. K., Arend, R., Rosenbaum, D., Binns, H., Dawson, N., Sobel, H., & Isaacs, C. (1998). Prevalence rates and correlates of psychiatric disorders among preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 204-214.
- Lavigne, J. V., Gibbons, R. D., Christoffel, K. K., Arend, R., Rosenbaum, D., Binns, H., Dawson, N., Sobel, H., & Isaacs, C. (1996). Prevalence rates and correlates of psychiatric disorders among preschool children. In M. E. Hertzig, & E. A. Farber, (Eds.), *Annual*

progress in child psychiatry and child development. (pp. 303-318). Bristol, PA:
 Brunner/Mazel, Inc.

- Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology, 28*, 129-137.
- Levine, F. M., & Fasnacht, G. (1974). Token rewards may lead to token learning. *American Psychologist, 29*, 817-820.
- Levine, F. M., & Fasnacht, G. (1976). "Token rewards may lead to token learning": Reply. *American Psychologist, 31*, 90-92.
- Margalit, M. (1983). Diagnostic application of the Conners' Abbreviated Symptom Questionnaire. *Journal of Clinical Child Psychology, 12*, 355-357.
- Martin, G., & Pear, J. (1996). *Behavior Modification: What it is and how to do it*. Saddle River, NJ: Prentice Hall.
- McConaughy, S. H., Kay, P. J., & Fitzgerald, M. (1999). The Achieving, Behaving, Caring project for preventing ED: Two-year outcome. *Journal of Emotional and Behavioral Disorders, 7*, 224-239.
- McGoey, K. E., & DuPaul, G. J. (2000). Token reinforcement and response cost procedures: Reducing the disruptive behavior of preschool children with Attention-Deficit/Hyperactivity Disorder. *School Psychology Quarterly, 15*, 330-343.
- McIntosh, D. E., Rizza, M. G., & Bliss, L. (2000). Implementing empirically supported interventions: Teacher-Child Interaction Therapy. *Psychology in the Schools, 37*, 453-462.

- McNeil, C. B. (2000). *The Tough Class Discipline Kit*. Congmont, CO: Sopris West.
- McNeil, C. B. (1995). *The AD/HD Classroom Kit: An inclusive approach to behavior management*. King of Prussia, PA: Center for Applied Psychology.
- McNeil, C. B., Capage, L. C., Bahl, A., & Blanc, H. (1999). Importance of early intervention for disruptive behavior problems: Comparison of treatment and wait-list control groups. *Early Education and Development, 10*, 445-454.
- McNeil, C. B., Eyberg, S., Eisenstadt, T. H., Newcomb, K., & Funderburk, B. (1991). Parent-Child Interaction Therapy with behavior problem children: Generalization of treatment effects to the school setting. *Journal of Clinical Child Psychology, 20*, 140-151.
- McNeil, C. B., & Filcheck, H. A. (in press). *The Level System*. Longmont, CO: Sopris West.
- Merrett, F., & Wheldall, K. (1993). How do teachers learn to manage classroom behaviour? A study of teachers' opinions about their initial training with special reference to classroom behaviour management. *Educational Studies, 19*, 91-106.
- Miller, M. A., McCullough, C. S., & Ulman, J. D. (1981). Carryover effects of multielement manipulations: Enhancement of preschoolers' appropriate rest-time behavior. *Educational Psychology, 1*, 341-346.
- Miller, L. K., & Schneider, R. (1970). The use of a token system in project Head Start. *Journal of Applied Behavior Analysis, 3*(3), 191-197.
- Miller, M. A., McCullough, C. S., & Ulman, J. D. (1981). Carryover effects of multielement manipulations: Enhancement of preschoolers' appropriate rest-time behavior. *Educational Psychology, 1*, 341-346.
- Miltenberger, R. G. (2000). *Behavior modification: Principles and procedures*. (2nd ed.). Pacific Grove, CA: Brooks/Cole Publishing Co.

- Molloy, G. N. (1979). Reinforcement effects on intrinsic interest. *Mental Retardation Bulletin*, 7, 31-44.
- Murry, M. E. (1980). Behavioral management of the hyperactive child. *Journal of Developmental and Behavioral Pediatrics*, 1, 109-111.
- Musgrove, W. J. (1981). The establishment and decline of a token reinforcement program with three emotionally disturbed pre-school children. *Educational Psychology*, 1, 81-85.
- Odom, S. L., Hoyson, M., Jamieson, B., & Strain, P. S. (1985). Increasing handicapped preschoolers' peer social interactions: Cross-setting and component analysis. *Journal of Applied Behavior Analysis*, 18(1), 3-16.
- Okovita, H. W., & Bucher, B. (1976). Attending behaviour of children near a child who is reinforced for attending. *Psychology in the Schools*, 13, 205-211.
- O'Leary, K. D., & O'Leary, S. G. (1977). *Classroom management: The successful use of behavior modification*. New York: Pergamon Press Inc.
- O'Leary, K. D., Poulos, R. W., & Devine, V. T. (1972). Tangible reinforcers: Bonuses or bribes? *Journal of Consulting and Clinical Psychology*, 38, 1-8.
- Parsonson, B. S., & Baer, D. M. (1986). The graphic analysis of data. In A. Poling & R. W. Fuqua (Eds.), *Research methods in applied behavior analysis: Issues and advances* (157-186). New York: Plenum Press.
- Polyson, J. A. (1979). Toward a better view of children's problem behavior. *Psychology: A Journal of Human Behavior*, 16, 33-37.
- Reitman, D. (1998). The real and imagined harmful effects of rewards: Implications for clinical practice. *Journal of Behavior Therapy and Experimental Psychiatry*, 29, 101-113.

- Reynolds, L. D., & Kelley, M. L. (1997). The efficacy of a response const-based treatment package for managing aggressive behavior in preschoolers. *Behavior Modification, 21*, 216-230.
- Roane, H. S., Fisher, W. W., & McDonough, E. M. (2003). Progressing from programmatic to discovery research: A case example with the overjustification effect. *Journal of Applied Behavior Analysis, 36*(1), 35-46.
- Rowbury, T. G., Baer, A. M., & Baer, D. M. (1976). Interactions between teacher guidance and contingent access to play in developing preacademic skills of deviant preschool children. *Journal of Applied Behavior Analysis, 9*(1), 85-104.
- Skinner, C. H., Cashwell, C. S., & Dunn, M. S. (1996). Independent and interdependent group contingencies: Smoothing the rough waters. *Special Services in the Schools, 12*, 61-78.
- Storey, K., Danko, C. D., Ashworth, R., & Strain, P. S. (1994). Generalization of social skills intervention for preschoolers with social delays. *Education and Treatment of Children, 17*, 29-51.
- Swiezy, N. B., Matson, J. L., & Box, P. (1992). The Good Behavior Game: A token reinforcement system for preschoolers. *Child and Family Behavior Therapy, 14*(3), 21-32.
- Titus, S., Savage, M. A., Krebs, F., Aquino, D., Simonet, M., & Sachs, J. (1990). The process of changing from a level system to a token economy. *Residential Treatment for Children and Youth, 7*, 75-84.
- Turnbull, J. (1988). Perils (hidden and not so hidden) for the token economy. *Journal of Head Trauma Rehabilitation, 3*(3), 46-52.

- Vasta, R., & Stirpe, L. A. (1979). Reinforcement's effects on three measures of children's interest in math. *Behavior Modification, 3*, 223-244.
- Wainwright, A., & MHS Staff. (Eds.) (1996). *The Conners' Rating Scales: Over 25 years of research: An annotated bibliography*. North Tonawanda, New York: Multi-Health Systems Inc.
- Walker, H. M. (1983). Application of response cost in school settings: Outcomes, issues and recommendations. *Exceptional Education Quarterly, 3*, 46-55.
- Witt, J. C., Elliott, S. N., & Martens, B. K. (1984). Acceptability of behavioral interventions used in classrooms: The influence of amount of teacher time, severity of behavior problems, and type of intervention. *Behavioral Disorders, 9*, 95-104.
- Witt, J. C., & Martens, B. K. (1983). Assessing the acceptability of behavioral interventions used in classrooms. *Psychology in the Schools, 20*, 510-517.
- Wolfe, V. V., Boyd, L. A., & Wolfe, D. A. (1983). Teaching cooperative play to behavior problem preschool children. *Education and Treatment of Children, 6*, 1-9.

Appendix A

The Level System Manual

The Level System will be used as a way to provide positive attention to appropriate behavior and give a warning and a minor consequence for inappropriate behavior.

To implement Level System for the whole class.

- Each child is assigned a certain shape on the System.
- The Level System should be placed somewhere in the room so that the children are able to see where their shapes are on the System.
- It should remain visible during the entire class period.
- The Level System only will be used during the 9:20 to 10:20 and 10:20 to 11:20 class periods (and the transition between the class periods) with the children who are involved in the research project.

Moving Up

- Children move up for appropriate behavior (e.g., sitting correctly, putting coat in the cubby, following class rules) and down for inappropriate behavior (i.e., annoying and obnoxious behavior, not following class rules).
- When children move up, they are given a labeled praise (i.e., specific praise such as “Thank you for sitting in your seat”) for the appropriate behavior.

Moving Down

- Children first are given a warning for inappropriate behavior, and then if they do not begin to behave appropriately, they move down. For example, the teacher would say, “You have two choices. You can either put the crayons in the basket, or you will move down into the cloudy area.” Alternatively, the teacher can use the visual two choice warning signal without words and simply say the child’s name. Then, if the child did not begin to behave appropriately, the teacher would move the child down.
- Children are not given a warning for hurting others (e.g., hitting other children and making them cry) or destruction of property (e.g., tearing up other children’s work); they simply are moved down a level. The teacher has the option of adding a consequence (e.g., time out) when the child is moved down for these behaviors.
- If a child continues to misbehave after being moved down a level, give the child another warning if appropriate (i.e., not hurting or destruction), and if he or she still continues to misbehave, move the child’s shape down a level again. If the misbehavior continues after you have moved the child’s shape down 2 times, provide another consequence, such as time out, that you would typically give to a child who engaged in that misbehavior.

Rewards

- Near the end of the one-hour class period, the teacher should give out the rewards. All children in the sunny area of the Level System receive a reward (e.g., snack, activity), and all children in the cloudy area do not receive the reward.
- The rewards are printed on cards, and all of the children receive the same reward. The teacher should give out the rewards at least once during the one-hour class period.
- Each reward card should be used once before any reward card is used again. Place the used reward cards in the envelope so that each teacher knows which cards have been used.
- After the reward is given to the children, all of the children's shapes are placed back in the neutral area, and a new period begins where they can earn a reward. The children essentially are starting over for the next period.
- All of the children in the class should have equal access to the rewards. To ensure that the children are receiving the rewards with the same frequency:
 - ✓ Expectations must be individualized for each child, so that some children will move up for simply not hurting another child for a few minutes, or for staying seated for a few seconds. These expectations should be increased when the child masters them.
 - ✓ The teacher must complete the Daily Reward Log to monitor which children receive a reward or not. If some children continually are not receiving the reward, then that child's expectations must be lowered so that he or she may have access to the reward.
- The Level System should be used for all transitions (e.g., moving from circle time to structured activity, play time to clean-up). In other words, the teacher should move all of the names (either up or down) following every transition.

What if a child is in the 3rd cloudy level?

If a child is in the 3rd cloudy level of the system, take the child aside and discuss the rules with that child. In addition, the teacher should lower his or her expectations for that child and focus on small positive behaviors so that the child will receive positive feedback and be able to move up the Level System. For example, if a child is behaving aggressively, the teacher can tell the child that if he or she can play gently, the shape will be moved up a level. Another alternative would be to provide a reward to the class quickly and start all children back in the neutral zone.

What if a child is in the 3rd sunny level?

If a child is in the 3rd sunny level of the system, the child should still receive a labeled praise for behaving appropriately even though he/she cannot move up any more.

Correct Use of the Level System

1. Moving children's shapes up a level for appropriate behavior.
2. Providing a warning for inappropriate behavior.

3. Providing a warning, and then moving the child's shape down when the inappropriate behavior continues.
4. Providing a warning, then moving the child's shape down, then providing a warning again, when moving the child's shape down again when the inappropriate behavior continues.
5. Providing a warning, then moving the child's shape down, then providing a warning again, then moving the child's shape down again, then providing a back up consequence (e.g., time out) when the inappropriate behavior continues.

Common Problems with the Level System

1. No warning is given for inappropriate behavior (not hurting or destruction of property), the child's shape simply is moved down.
2. Providing a warning in a critical manner.
3. Forgetting to label praise the child's behavior when the child's shape is moved up.
4. Providing a warning, but forgetting to move the child's shape down when the behavior continues.
5. Using criticism or nagging instead of providing a warning for misbehavior.
6. Providing a warning for hurting or destruction of property.
7. Not praising children's behavior or moving them up for appropriate behavior.
8. Providing more than 1 warning before moving the child's shape down.
9. Not providing a warning or moving a child's shape down for inappropriate behavior.

Appendix B

Level System Rewards

#1 STICKER**#2 TREAT****#3 HAND STAMP****#4 CHARADES**

Rules: Up to three children will simultaneously act out one of these charades while the other players try to guess. No words are allowed.

- | | | |
|--------------------------|----------------------------|-----------------------------|
| • Digging | • Frog | • Standing on your head |
| • Cutting with scissors | • Monkey | • Opening a door |
| • Sleeping | • Shooting a bow and arrow | • Brushing your hair |
| • Laughing | • Fighting | • Juggling |
| • Eating corn on the cob | • Swimming | • Flying |
| • Kicking a ball | • Sneezing | • Standing up |
| • Monster | • Cat | • Standing on your tip toes |
| • Dancing | • Playing the piano | • Jumping |
| • Cooking | • Skipping | • Playing soccer |
| • Skating | • Blowing bubbles | • Catching a ball |
| • Taking a bath | • Stretching | • Holding your breath |
| • Flying a kite | • Playing the drums | • Throwing a snowball |
| • Falling down | • Driving a car | • Playing cards |
| • Waking up | • Reading a book | • Playing baseball |

#5 ACTIVITY

Rules: Teacher chooses one of the following:

- Act like a tightrope walker in the circus.
- Get a partner and pretend like you are taking a dog for a walk; one player will be the owner and the other will be the dog. Then switch.
- Act like a monkey eating a banana.
- Players stand in lines with legs apart. Children take turns being a snake and crawling through the tunnel.
- Choose a partner. Pretend like you are a dog fetching a stick. One player pretends to be the dog's owner. The other pretends to be the dog. Then switch.
- Act like you are building a sand castle at the beach.
- Act like your favorite zoo animal.
- Walk across the room holding your ankles.
- Close your eyes. Try to describe the clothes that the child next to you is wearing.
- Pretend to be a juggling clown.
- Act like a baby who is just learning to walk.
- With a partner, try to hold hands and walk backwards across the room.
- Pretend to be an airplane taking off.
- Act like a cheerleader at a football game.
- Do the bear walk. Walk with your hands and feet (but not your knees) on the floor.
- Act like an angry tiger.
- Walk across the room using baby steps (touching your heel to the tips of your toes).
- See how far you can walk with a piece of paper (or a book) on your head.
- Pretend to be the slowest turtle in the world.
- Wiggle your body all over while singing "Twinkle, Twinkle, Little Star."

#6 GUESS WHO'S MISSING

Rules: Children sit randomly on the floor. Everyone will put their heads in their lap and close their eyes. The teacher will tap someone on the head and ask him or her to hide (e.g., behind the teacher's desk, in the hall). Children will try to figure out who is missing. They are to keep quiet until everyone has had a chance to think it over. Then they can guess out loud. (As a variation, teachers can choose two children to hide.)

#7 SIMON SAYS

Rules: Children are to follow the leader's instructions only if the words "Simon Says" follow the instruction. Either the teacher or a chosen student can be the leader. If a child engages in the instruction without the "Simon Says" preface, the child must sit out. Examples of "Simon Says" instructions include: "put your hands on your knees," "touch your nose," "turn around in a circle," "shake your neighbor's hand," and "give yourself a hug."

#8 PROHIBITED MOVEMENTS

Rules: The children form a circle. They stand at the distance of stretched arms from each other. The teacher chooses a leader to be in the center of the circle. He/she makes different movements and the other children repeat all of them exactly. But there are 2 movements that the children mustn't repeat. Instead of them they must make other movements. The teacher tells the children what these movements are at the beginning of the game (e.g., jumping, clapping hands, bending knees). The children who make mistakes are out.

#9 LIMBO CONTEST

Rules: Teacher holds out a pointer (ruler or stick). Children try to go under the pointer by leaning backwards. Children who touch the pointer or fall down are eliminated.

#10 HOT POTATO

Rules: Have the players sit in a circle. Use a bean bag (teddy bear, book) as the "hot potato". Choose one child to sit with his or her back to the group. Once the children begin passing the "hot potato" around quickly, the child will shout, "HOT POTATO!". Whoever has the potato at that time will have to sit out.

#11 FOLLOW THE LEADER

Rules: Teacher will choose a leader. All of the other players will line up behind the leader. The leader will be instructed to walk around and do funny things (like crawling, jumping, patting the desk, whistling). Sequentially (taking turns down the line), each person will follow the leader's activities.

#12 BAG (BALL) TOSS

Rules: Teacher cleans out a trash can (box or bag). Using a bean bag or ball, have each child try to throw the object into the can. The game can be varied by having the children stand further away or having them bounce the ball before it enters the can.

#13 OVER AND UNDER RELAY

Rules: Have players line up in two lines. Using any object (block, stuffed animal, bean bag), have players hand the object to the person behind them. The first player hands it over their head. The second player accepts it up high but then passes it through the legs. The passing is alternated, over then under then over then under. When they get to the end of the line, they turn around and go back the other way. The first team to relay the object back to the line leader wins.

#14 "STOP THE MUSIC! I'M TIRED."

Rules: Using music or hand clapping, have players walk around in a circle. When music or clapping stops, the children are to sit on their bottoms as quickly as possible. The last child to sit on his or her bottom must sit out. Then the music/clapping begins again. (If chairs are used, this game is called "musical chairs".) Children can be fooled by changing the volume or pace. If a child sits before the music/clapping stops, (s)he must sit out.

#15 PASS A HAT

Rules: The children are in a circles. One child has a hat on his/her head. The object is to pass the hat round the circle. But the children mustn't use their hands while passing the hat from head to head.

#16 HOT POTATO/NAME TOSS

Rules: Choose a "hot potato" caller. Then, have the rest of the children stand in two lines that face each other. Using a bean bag, ball, or stuffed animal as the "hot potato", have children throw the "hot potato" to each other as quickly as possible. The trick is that they have to say the person's name (can also use last names and initials) before throwing the "hot potato." Whoever is holding the hot potato when the caller yells "HOT POTATO!" has to sit out.

#17 DUCK, DUCK, GOOSE

Rules: Teacher chooses one person to be "it" and walk around the circle. As they walk around, they tap people's heads and say whether they are a "duck" or a "goose". Once someone is the "goose" they get up and try to chase "it" around the circle. The goal is to tap that person before they are able sit down in the "goose's" spot. If the goose is not able to do this, they become "it" for the next round and play continues. If they do tap the "it" person, the person tagged has to sit in the center of the circle. Then the goose becomes it for the next round. The person in the middle can't leave until another person is tagged and they are replaced.

#18 RED LIGHT, GREEN LIGHT

Rules: All the kids line up away from the teacher. When the teacher says green light, all the kids run, walk, crawl, hop (teacher's decision) as fast as they can. When the teacher says red light, everyone stops. Whoever doesn't stop is out. The teacher keeps saying red light or green light until all the kids are out or has gone from one end of the room to the other.

#19 WHO HAS A TOY?

Rules: The teacher puts toys on the floor and has the children form a circle around the toys. The teacher then plays music. As the music plays everyone walks around the circle. When the music stops everyone must grasp a toy. A person who doesn't have a toy is out. The number of toys should lessen every time.

#20 DOGGY, DOGGY, WHERE'S YOUR BONE?

Rules: A student is chosen to be the "dog." The dog sits with his or her back to the class and eyes closed. An eraser or another object (bone) is put under the chair. The teacher chooses someone to sneak up and steal the bone and hide it somewhere on themselves. Then everyone sings: "Doggy, Doggy, where's your bone? Somebody's stole it from your home. Guess who it might be." Then the dog has three chances to guess who took it. If the dog guessed right then he got to do it again. If he guessed wrong than the person who had the bone got a turn as the dog.

#21 BUBBLE GUM

Rules: Everyone sits in a tight circle and makes a fist with both hands and holds their fists out in the center of the circle. The teacher chooses a counter who takes one of his fists and taps everyone's fists (including his own) as he says this rhyme: "Bubble gum, Bubble gum, in a dish. How many pieces do you wish?" Who ever he taps last would give a number between 1 and 10. Then the counter taps everyone's fists as he/she counts to the specified number. Who ever he landed on had to take that fist out of the circle and place their hand behind their back. This can continue until the time is up.

#22 TOSS THE BALL

Rules: The children form a circle. The teacher chooses one person to start with the ball. He/she names a particular person in the group and throws the ball to him/her. That person must catch the ball then names another person of the group and also throws the ball to him/her.

#23 SAUSAGE

Rules: Everyone sits in a circle. The person who the teacher chooses to be "it" stands in the center of the circle. Each person asks the person who is "it" an appropriate question. The only answer to every question can be sausage. The first person to make the center person laugh wins a try in the middle. Ex: What color is your hair? Sausage. What do you brush your teeth with? Sausage.

#24 BUTTON, BUTTON, WHO HAS THE BUTTON?

Rules: The children all stand in a circle with one person in the middle. That person has a button between his hands that he is going to drop into one of the other player's hands. The object of the game is to go around the circle pretending to drop the button, but not letting the others know who has the button. Then the other players guess who has the button. Whoever finally guesses correctly gets to be in the middle and drop the button.

#25 WINDOWS AND DOORS

Rules: The children form a circle holding hands. Then they spread out enough so that everyone's arms are straight out, to form large spaces between kids. These are the windows and doors. Then the teacher chooses one child to start running, and weaving in and out between children. The children in the circle randomly drop their arms down trying to touch or trap the person weaving their way in and out. Once the person is caught or touched by the arms of someone, they are out. The child that caught them is the next to weave in and out of the windows and doors.

#26 CHINESE SCHOOL

Rules: The teacher says "Chinese school has just begun. No more laughing, no more fun. If you show your teeth or tongue, you will have a penalty done." The teacher then precedes to do silly things to make the children laugh. The kids that laugh hard enough that their mouths opened, exposing teeth and/or tongues, have to do whatever the teacher tells them to do (run around the room, hop on one leg, do jumping jacks).

#27 SLEEPING LIONS

Rules: All of the children (except one or two hunters) lie down on the floor in sleeping positions. Once they are settled, they are not allowed to move. The hunters walk through the room and try to make the sleeping lions move by making them laugh, and telling them jokes. The hunters are not allowed to touch the lions. Once the lions have moved, they get up and join the hunters.

#28 LION'S CUB

Rules: The teacher selects a person to be the Lion. The lion sit with their back to the other children at least ten feet away. The teacher puts a stuffed animal behind the lion and has the lion pretend it is his cub. The other children take turns sneaking up behind the Lion and trying to steal the cub. If the lion hears the person sneaking up it can roar and then turn around. If the lion has caught a child, that child takes the lions place and the lion goes back to the other children.

#29 LAND, WATER

Rules: The teacher says the words "land" and "water." On hearing the word "land" the children jump forward, onhearing the word "water" backward. When a child makes a mistake, her or she is out.

Appendix C

Conner's Global Index

Child Rated: _____ **Rater:** _____ **Date:** _____

Instructions

Read each item below carefully, and decide how much you think your student has been bothered by this problem in the past month. For each behavior described below, circle one number to indicate how much of a problem the behavior was for your student.

	Not at all	Just a little	Pretty Much	Very much
1. Restless in the "squirmy" sense	0	1	2	3
2. Temper outbursts, unpredictable behavior	0	1	2	3
2. Distractibility or attention span a problem	0	1	2	3
4. Disturbs other children	0	1	2	3
5. Pouts and sulks	0	1	2	3
6. Mood changes quickly and drastically	0	1	2	3
7. Restless, always up and on the go	0	1	2	3
8. Excitable, impulsive	0	1	2	3
9. Fails to finish things	0	1	2	3
10. Easily frustrated in efforts	0	1	2	3

Appendix D

Demographic Questionnaire

Please complete the following information. Place an X next to the appropriate responses.

Child's Name _____ Date _____

Child's Age _____

Child's Gender: ___ Male
 ___ Female

Child's Race: ___ Caucasian
 ___ African American
 ___ Bi-racial
 ___ Asian
 ___ Hispanic
 ___ Other _____

The child's primary caregivers are: ___ Mother only
 ___ Father only
 ___ Mother and Father
 ___ Mother and Stepfather/Significant Other
 ___ Father and Stepmother/Significant Other
 ___ Other _____

The child's primary caregivers are: ___ Married
 ___ Separated
 ___ Divorced
 ___ Single
 ___ Living Together
 ___ Widow/er

The following child's primary caregivers are employed:

___ Mother only	As what? _____
___ Father only	As what? _____
___ Mother and Father	As what? _____
	As what? _____
___ Mother and Stepfather/Significant Other	As what? _____
	As what? _____
___ Father and Stepmother/Significant Other	As what? _____
	As what? _____
___ Stepparent/Significant Other only	As what? _____
___ Other _____	As what? _____

The highest education level of the child's primary caregivers are:

Caregiver: _____

- Some High School
- GED
- High School Degree
- Professional/Trade School Degree
- Associate's Degree
- Bachelor's Degree
- Master's Degree
- Advanced Degree (ex: Ph.D., M.D.)
- Other _____

Caregiver: _____

- Some High School
- GED
- High School Degree
- Professional/Trade School Degree
- Associate's Degree
- Bachelor's Degree
- Master's Degree
- Advanced Degree (ex: Ph.D., M.D.)
- Other _____

Approximate family income per year:

- \$10,000 or below
- \$10,001 to \$20,000
- \$20,001 to \$30,000
- \$30,001 to \$40,000
- \$40,001 to \$50,000
- \$50,001 to \$60,000
- \$60,001 or above

Appendix E

Revised Edition of the School Observation Coding System

Definitions of Behavior

Category: Inappropriate Behavior

A. *Appropriate Behavior*: The absence of inappropriate behavior for the entire 10-s interval.

If unsure as to whether the behavior was appropriate or inappropriate, code Appropriate Behavior.

B. *Inappropriate Behavior*: Behaviors are coded as Inappropriate Behaviors because they are annoying or disruptive to the target child, the teacher, or other children.

Definitions of Inappropriate Behaviors:

1. *Whining* – Coherent words uttered by the child in a slurring, nasal, high-pitched, voice.
2. *Crying* – Inarticulate utterances of distress (e.g., audible weeping) that may or may not be accompanied by tears.
3. *Yelling* – Loud screeching, screaming, shouting, or crying. The sound must be loud enough so that it is clearly above the intensity of normal indoor conversation. Not coded during outdoor recess observations.
4. *Destructiveness* – Behaviors during which the child damages or destroys an object or threatens to damage an object. Do not code destructiveness if it is appropriate within the context of play situation (e.g., ramming cars in a car crash).
5. *Aggressive Behavior* – Examples include fighting, kicking, slapping, hitting, grabbing an object roughly from another person, or threatening to do any of the preceding.

6. *Negativism* – Verbal or nonverbal behavior expressing a negative attitude.

Negativism may be scored when the child makes a neutral comment that is delivered in a tone of voice that conveys an attitude of “don’t bother me.” Negativism may be expressed in a derogatory, uncomplimentary, or angry manner. Also included are defeatist statements such as “I give up”, contradictions of another person, and teasing or mocking behaviors or verbalizations. “Pouting” facial expressions are included in this category.

7. *Self-Stimulation* – Repetitive physical movements (involving only the child’s body and not other objects) that may be harmful and that interfere with a child’s ability to attend or complete a task. Examples include head-banging, thumb-sucking, and masturbation.
8. *Demanding Attention* – Includes inappropriate verbal or nonverbal requests for attention from the teacher or other students (e.g., “Call on me! Call on me! Call on me!). Examples include tugging on the teacher’s sleeve, tapping a neighbor on the shoulder, waving arms in the air, and passing notes to another child.
9. *Disruptive Behavior* – Any physically active or repetitive behavior that is or may become disruptive to others or interfere with the target child’s ability to attend or complete a task. Examples include kicking a child’s chair repeatedly, drumming on the table loudly, clowning, making funny noises, teasing, or spinning a pencil on the desk.
10. *Talking Out of Order* – Any talking when the class has been instructed to be silent unless called on to speak. This includes situations in which a “classroom rule” exists that silence is to be maintained (i.e., the teacher does not have to give the instruction

- explicitly – the expectation for silence is sufficient). Examples include whispering to a neighbor, answering a question directed to someone else, calling out to another child, and talking, singing, or humming to oneself.
11. *Being Out of Area* – Coded when the target child, without permission, leaves the area to which s/he is assigned. Examples include standing up when the rest of the class is seated, leaving the desk, approaching the teacher without permission, or playing with a toy that is not in the child's assigned work area the child is suppose to be. The behavior must be appropriate for the context or classroom norms (e.g., in some classrooms children are allowed walk to the teacher's desk to obtain help with an assignment).
- C. *Not Applicable*: Coded when there is no readily identifiable task that the child is expected to perform. Examples of Not Applicable activities include free play and unstructured recess time.

Appendix F

Behavioral Observation Coding Sheet

Child #1: _____ Teacher: _____ Date: _____

Child #2: _____ Time: _____ Coder: _____

Child #3: _____

Child #4: _____

Child # ____															
Interval	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Appropriate															
Inappropriate															
Response															
Interval	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Appropriate															
Inappropriate															
Response															

Response Key: LP = Labeled Praise; W = Warning; TO = Time Out; ↑ = Moved Up a Level;

↓ = Moved Down a Level; N = No Response

Inappropriate Behaviors: whining, crying, yelling, destructiveness, aggressive behavior, negativism, self-stimulation, demanding attention, disruptive behavior, talking out of order, being out of area

Behavioral Observation Coding Sheet

Child #1: _____ Teacher: _____ Date: _____

Child #2: _____ Time: _____ Coder: _____

Child #3: _____

Child #4: _____

Child # ____															
Interval	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Appropriate															
Inappropriate															
Response T															
TA															
Interval	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Appropriate															
Inappropriate															
Response T															
TA															

Response Key: LP = Labeled Praise; W = Warning; TO = Time Out; ↑ = Moved Up a Level;

↓ = Moved Down a Level; N = No Response; O = Other

Inappropriate Behaviors: whining, crying, yelling, destructiveness, aggressive behavior, negativism, self-stimulation, demanding attention, disruptive behavior, talking out of order, being out of area

Appendix G

Classroom Management Rating

Teacher's Name _____ Day _____ Date _____

How manageable was your class today?

1
Completely
Unmanageable

2

3
Somewhat
Unmanageable/
Manageable

4

5
Completely
Manageable

Appendix I

Intervention Rating Profile

Teacher Name: _____ Date: _____

Instructions

Read each item below carefully and circle one number to indicate how much you agree/disagree with the statement as it applies to the program circled above.

	Strongly Disagree					Strongly Agree
	1	2	3	4	5	6
1. Most teachers would find the intervention suitable for behavior problems in the classroom.	1	2	3	4	5	6
2. Most teachers would find this intervention appropriate for behavior problems in addition to those that I have experienced in the classroom.	1	2	3	4	5	6
3. Children’s behavior problems in the classroom are severe enough to warrant the use of this intervention.	1	2	3	4	5	6
4. This intervention should prove effective in changing children’s problem behavior.	1	2	3	4	5	6
5. This would be an acceptable intervention for children’s problem behavior.	1	2	3	4	5	6
6. Overall, the intervention would be beneficial for children.	1	2	3	4	5	6
7. I would be willing to use this intervention in the classroom setting.	1	2	3	4	5	6
8. This intervention would be appropriate for use <i>before</i> making a referral.	1	2	3	4	5	6
9. This intervention would not result in negative side effects for children.	1	2	3	4	5	6

	Strongly Disagree					Strongly Agree
	1	2	3	4	5	6
10. This intervention would not result in risk to children.	1	2	3	4	5	6
11. This intervention would not be considered a “last resort.”	1	2	3	4	5	6
12. This intervention is practical in the amount of time required for parent contact.	1	2	3	4	5	6
13. This intervention is practical in the amount of time required for contact with day care staff.	1	2	3	4	5	6
14. This intervention is practical in the amount of time required for record keeping.	1	2	3	4	5	6
15. This intervention is practical in the amount of out-of-day care time required for implementation.	1	2	3	4	5	6
16. This intervention would not be difficult to implement in a classroom with 30 students.	1	2	3	4	5	6
17. This intervention would not be disruptive to students.	1	2	3	4	5	6
18. It would not be difficult to use this intervention and still meet the needs of children in the classroom.	1	2	3	4	5	6
19. Teachers are likely to use this intervention because it requires little technical skill.	1	2	3	4	5	6
20. Teachers are likely to use this intervention because it requires little training to implement effectively.	1	2	3	4	5	6

Appendix J

Parent Interview

Parent’s Name: _____ Child’s Name: _____

Date: _____

Instructions

We are interested in how parents feel about different classroom management strategies. Read each item below carefully and circle one number to indicate how acceptable/unacceptable that you feel each strategy is to use in the classroom to manage behavior.

	Very Unacceptable					Very Acceptable
	1	2	3	4	5	6
1. The use of a Level System in which children move up for appropriate behavior and move down for inappropriate behavior. At certain times throughout the day, the children who are in the sunny zone receive a reward.	1	2	3	4	5	6
2. Your child does not receive the reward from the level system for 2 days in a row.	1	2	3	4	5	6
3. Posting the level system on the wall so that everyone who enters the room can see it.	1	2	3	4	5	6
4. Praising children for appropriate behavior. For example, “Good job sharing your toy!”	1	2	3	4	5	6
5. The use of “when-then” statements. For example, “ <i>When</i> you sit on your carpet, <i>then</i> I will read the book.”	1	2	3	4	5	6
6. Using redirection, such as moving a child to another area or getting them interested in another activity, when mild misbehavior occurs (like arguing).	1	2	3	4	5	6
7. Giving children “two-choices” with logical consequences for inappropriate behavior. For example, “You have <i>two-choices</i> . You can either play gently with the blocks, or I will put the blocks away.”	1	2	3	4	5	6

	Very Unacceptable				Very Acceptable	
	1	2	3	4	5	6
8. The use of a “calm down area” which is a place in the room, away from other children, where children go when they are becoming too hyper to take part in a quiet activity (like reading or doing a puzzle) until they calm down enough to return to the regular class activity.	1	2	3	4	5	6
9. The use of time-out in a chair in the room for dangerous or destructive behavior or repeated noncompliance for a maximum amount of time of 3 minutes. Children are given a “two-choices” warning (see above) before they go to time-out so that they can correct their misbehavior.	1	2	3	4	5	6
10. The current classroom management strategies that the teacher and teacher’s aide are using.	1	2	3	4	5	6
11. Please list all of the positive things about the level system.						
12. Please list all of the negative things about the level system.						
13. Please list all of the positive things about time-out.						
14. Please list all of the negative things about time-out.						
15. How do you feel about modifying expectations for individual children while using the level system so that children with behavior problems receive the reward about the same amount that the other children do? For example, most children in the classroom may be expected to sit on their carpet during reading time, and they may not move up a level. However, if a child has trouble sitting on his/her carpet, that child may be rewarded (move up a level) for sitting on his/her carpet until the child is able to sit just like the other children.						

Appendix K

Definitions

- A. *Labeled Praise*: A labeled praise lets the child know exactly what is liked about his or her behavior because it is specific in that it describes the behavior being praised (e.g., Nice job sharing the toys).

Appendix L

The Level System Treatment Integrity Checklist

Teacher: _____ Coder: _____ Date: _____ Time: _____

Question	Response for Teacher			Response for Teacher's Aide		
1. The Level System is hung in a location visible to all children in the classroom.	Yes	No	N/A	Yes	No	N/A
2. Each child has a shape with his or her name on it.	Yes	No	N/A	Yes	No	N/A
3. Most children move up a level for generally appropriate behavior at least two times within an hour.	Yes	No	N/A	Yes	No	N/A
4. Labeled praise is used when moving the child up a level most of the time.	Yes	No	N/A	Yes	No	N/A
5. Ten labeled praises given during the one-hour period that the Level System is used.	Yes	No	N/A	Yes	No	N/A
6. Teacher used a warning signal for the majority of inappropriate behaviors in the classroom.	Yes	No	N/A	Yes	No	N/A
7. Teacher followed through with moving the child down a level when a warning signal was given and the behavior continued most of the time.	Yes	No	N/A	Yes	No	N/A
8. A boring, monotonous tone of voice is used when moving the child down a level most of the time.	Yes	No	N/A	Yes	No	N/A
9. Children are given a warning before they move down a level every time (except for hurting).	Yes	No	N/A	Yes	No	N/A
10. No warning given for hurting or destruction of property. Child simply is moved down a level most of the time.	Yes	No	N/A	Yes	No	N/A
11. Children who are involved in the reward are able to move to a different or separate part of the room to participate.	Yes	No	N/A	Yes	No	N/A
12. The Level System is used for each major transition.	Yes	No	N/A	Yes	No	N/A

Appendix M
Daily Reward Log

Teacher's Name _____ Day _____ Date _____

Cross out the names of the children who did NOT receive the reward in your class today.

(Each child's name in the class was listed below.)

Table 1

T-Scores of Participants on CGI by Teacher per Condition

Participant and condition	Teacher					
	Patricia	Colleen	Lori	Chad	Rachel	All ^a
Luke						
Selection	69	69	71	70	73	70.4
Baseline	66	64	63	70	59	64.4
Level System	53	53	64	69	60	59.8
Withdrawal	66	66	66	70	56	64.8
Level System	64	67	64	70	58	64.4
Follow up	64	69	62	70	59	64.8
Cody						
Selection	63	52	70	64	53	60.4
Baseline	50	50	61	62	62	57.0
Level System	55	55	60	62	60	58.4
Withdrawal	64	58	66	64	55	61.4
Level System	68	50	59	67	54	59.6
Follow up	44	50	56	66	52	53.6
Tim						
Selection	44	46	47	46	43	45.2
Baseline	45	45	45	46	46	45.4
Level System	44	44	44	45	46	44.6
Withdrawal	44	45	42	45	43	43.8

Participant and condition	Teacher					
	Patricia	Colleen	Lori	Chad	Rachel	All ^a
Tim						
Level System	44	47	45	45	42	42.4
Follow up	44	46	46	47	42	45.0
John						
Selection	46	42	47	47	43	45.2
Baseline	42	42	42	43	42	42.2
Level System	42	42	42	42	42	42.0
Withdrawal	42	42	42	42	42	42.0
Level System	42	42	43	43	42	42.4
Follow up	42	42	43	43	42	42.2

Note. ^aThe values represent mean T-scores of teacher responses combined.

Table 2

Class Management Ratings by Teacher per Condition

Condition	Teacher					
	Patricia	Colleen	Lori	Chad	Rachel	All ^a
Baseline	4.8	4.3	4.3	3.8	4.4	4.3
Level System	4.7	4.7	4.3	4.3	4.3	4.4
Withdrawal	5.0	4.8	4.1	3.9	4.7	4.5
Level System	5.0	5.0	5.0	3.8	4.9	4.8
Follow up	5.0	4.7	4.4	3.7	5.0	4.6

Note. ^aThe values represent mean class management ratings of teacher responses combined.

Table 3

Mean Number of Time Outs Provided to Participants by Teacher per Condition

Participant and condition	Teacher					
	Patricia	Colleen	Lori	Chad	Rachel	All ^a
Luke						
Baseline	0.44	0	0.78	0.44	0.57	0.76
Level System	0	0.17	0.50	0.17	0.33	0.39
Withdrawal	0.14	0.13	0.44	0.56	0.29	0.54
Level System	0	0	0.17	0.50	0	0.20
Follow up	0	0	0	0.17	0	0.12
Cody						
Baseline	0.11	0	0.44	0.33	0	0.36
Level System	0	0.17	0.33	0	0	0.17
Withdrawal	0.43	0	0.56	0.33	0	0.46
Level System	0	0	0	0.33	0	0.10
Follow up	0	0	0	0.33	0	0.12
All children in class						
Baseline	0.14	0	0.31	0.19	0.14	2.04
Level System	0	0.08	0.21	0.04	0.08	1.06
Withdrawal	0.14	0.03	0.25	0.22	0.07	1.50
Level System	0	0	0.04	0.21	0	0.35
Follow up	0	0	0	0.13	0	0.24

Note. ^aThe values represent mean number of time outs provided by teachers combined.

Table 4

Percentage of Integrity Behavior Exhibited by Teachers in Response to Child Behavior by

Condition

Condition	Praise	Move shape up	Warning	Move shape down	Criticism	Time out	Other
Child appropriate behavior							
Baseline	5.8	0	0	0	3.3	0	57.8
Level System	18.9	7.3	0	0	.3	0	46.1
Withdrawal	6.4	0	0	0	2.7	0	64.6
Level System	15.8	6.9	0	0	.5	0	29.1
Follow up	5.8	0	0	0	.3	0	22.5
Child inappropriate behavior							
Baseline	2.6	0	0	0	17.8	3	66.7
Level System	3.6	2.9	26.2	8.5	5.6	1.6	38.8
Withdrawal	1.2	0	0	0	7.9	3.5	61.8
Level System	2.5	0	43	42	12.6	0	52.8
Follow up	0	0	0	0	20.2	1	29.3

Note. The values represent mean percentages of teacher behavior.

Figure 1. Percentage of intervals scored with appropriate behavior exhibited by each participant across teachers with horizontal lines indicating means for each phase.

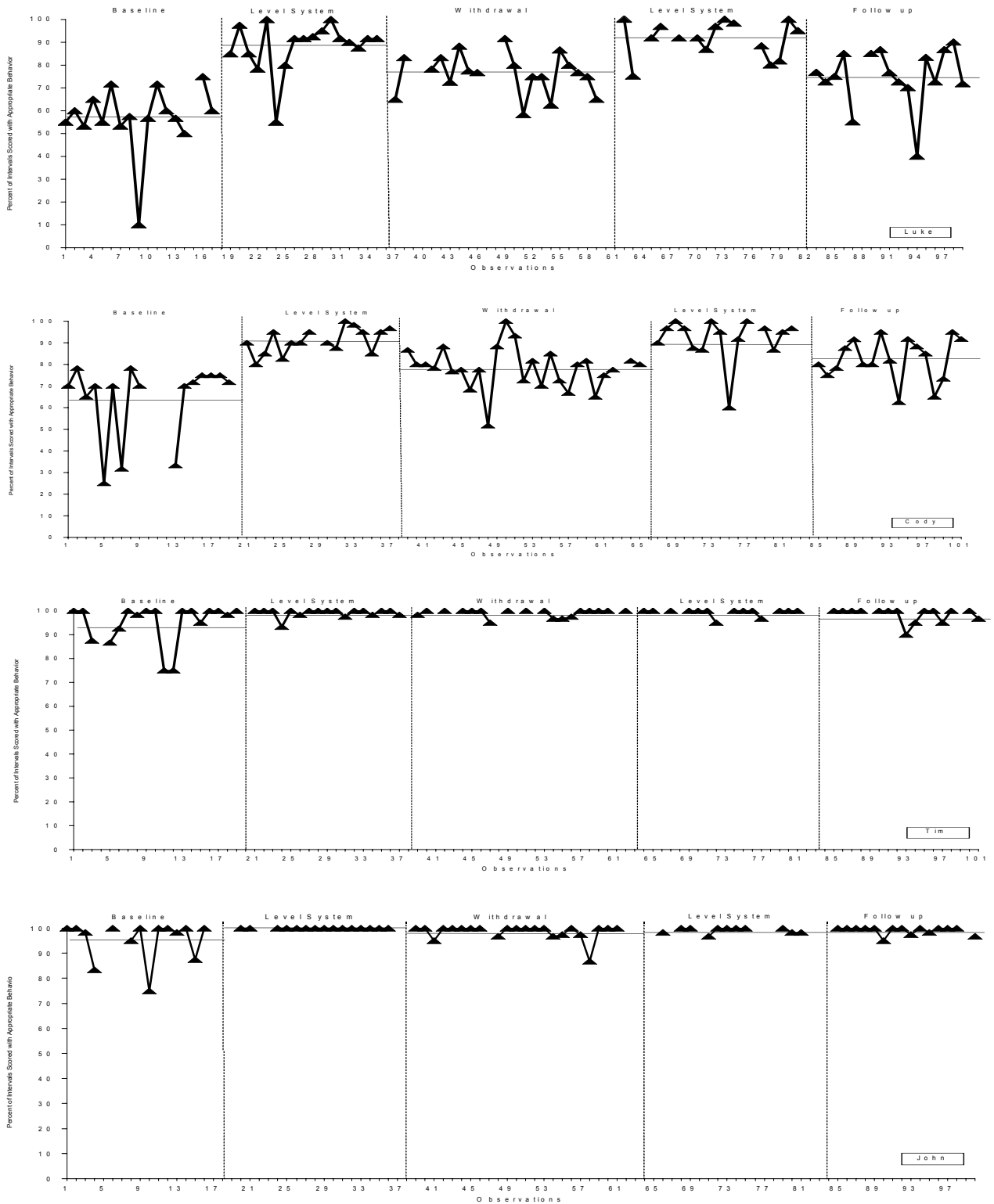


Figure 2. Percentage of intervals scored with appropriate behavior exhibited by Luke in each teacher's class with horizontal lines indicating means for each phase.

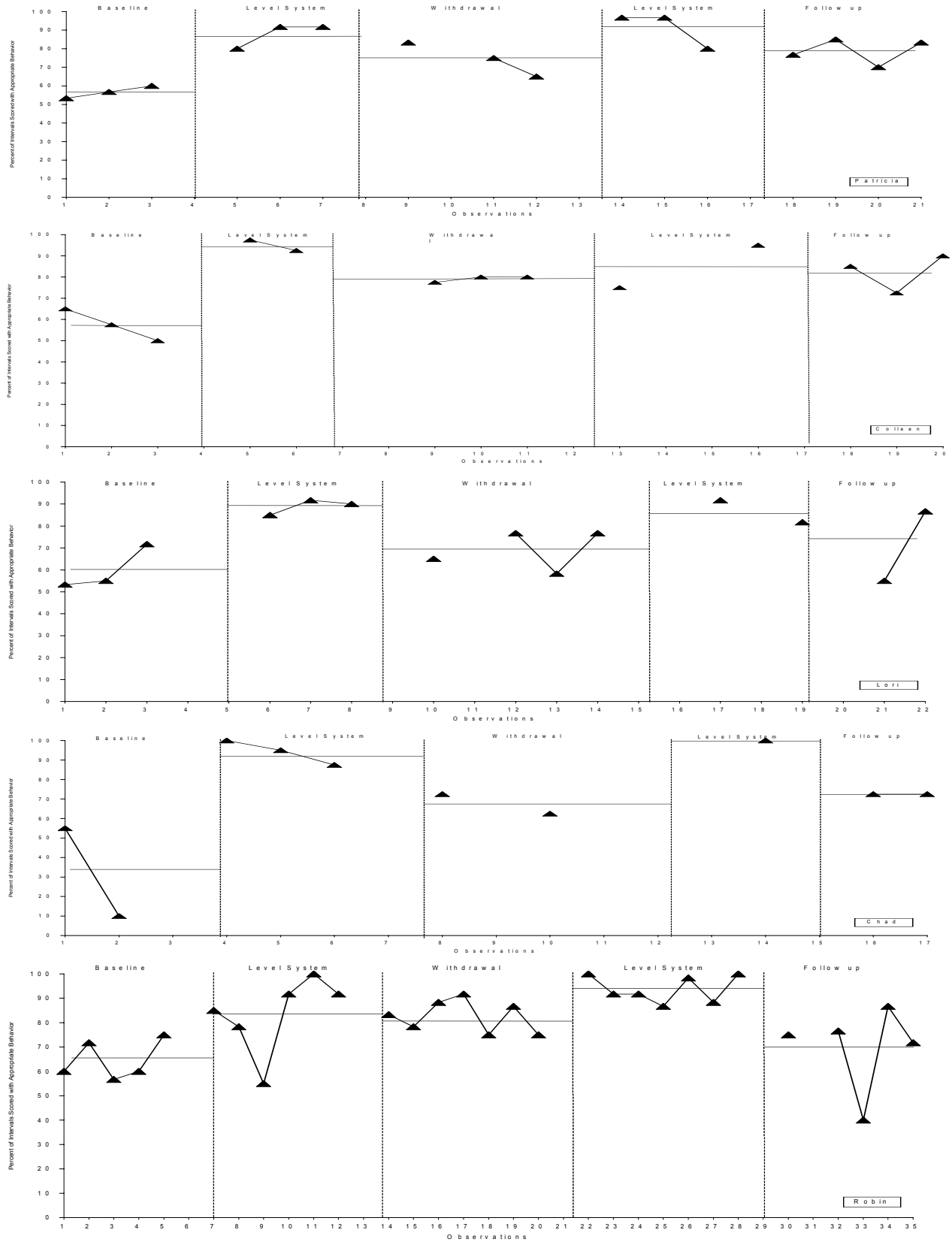


Figure 3. Percentage of intervals scored with appropriate behavior exhibited by Cody in each teacher's class with horizontal lines indicating means for each phase.

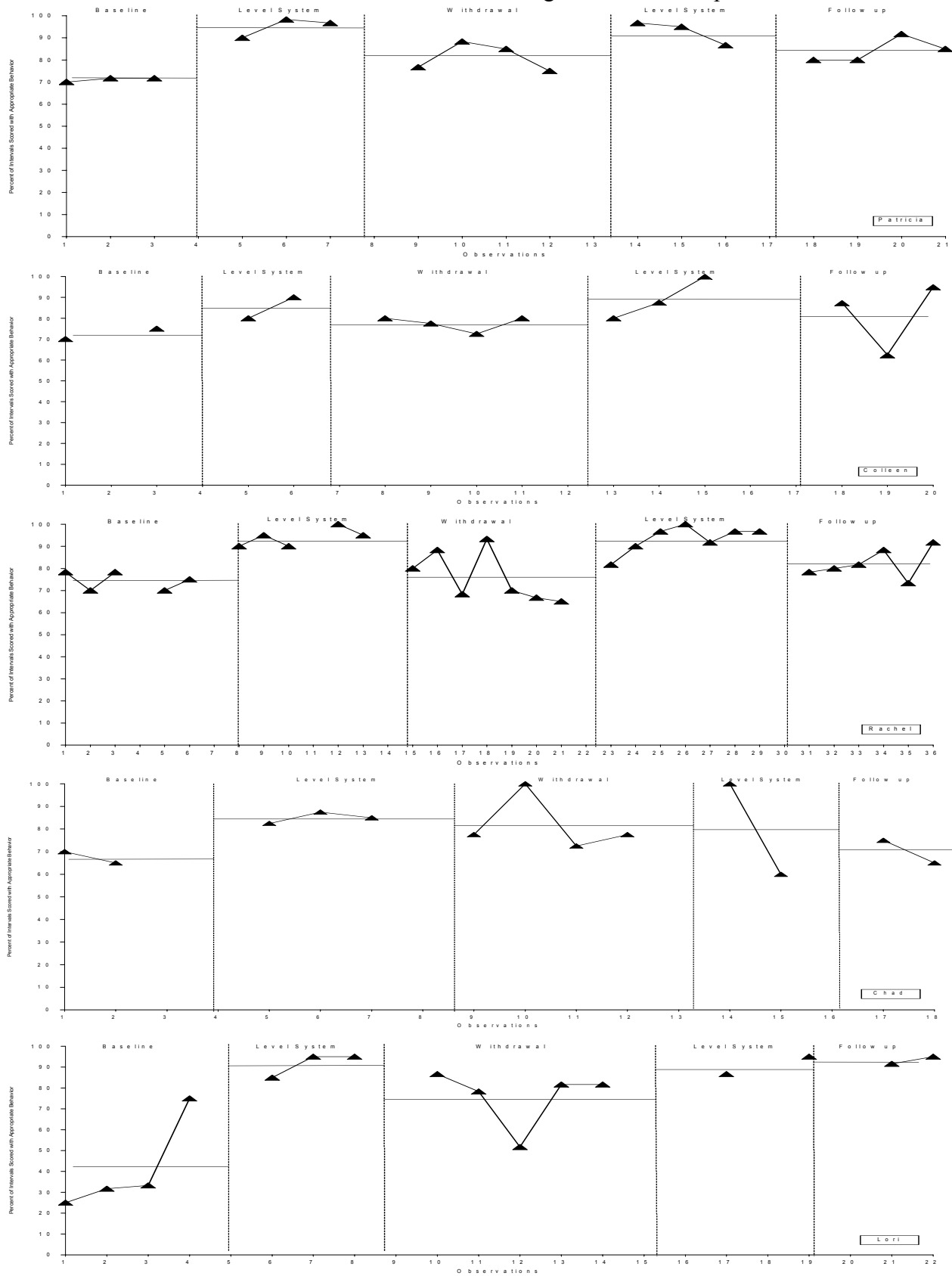


Figure 4. Percentage of intervals scored with appropriate behavior exhibited by Tim in each teacher's class with horizontal lines indicating means for each phase.

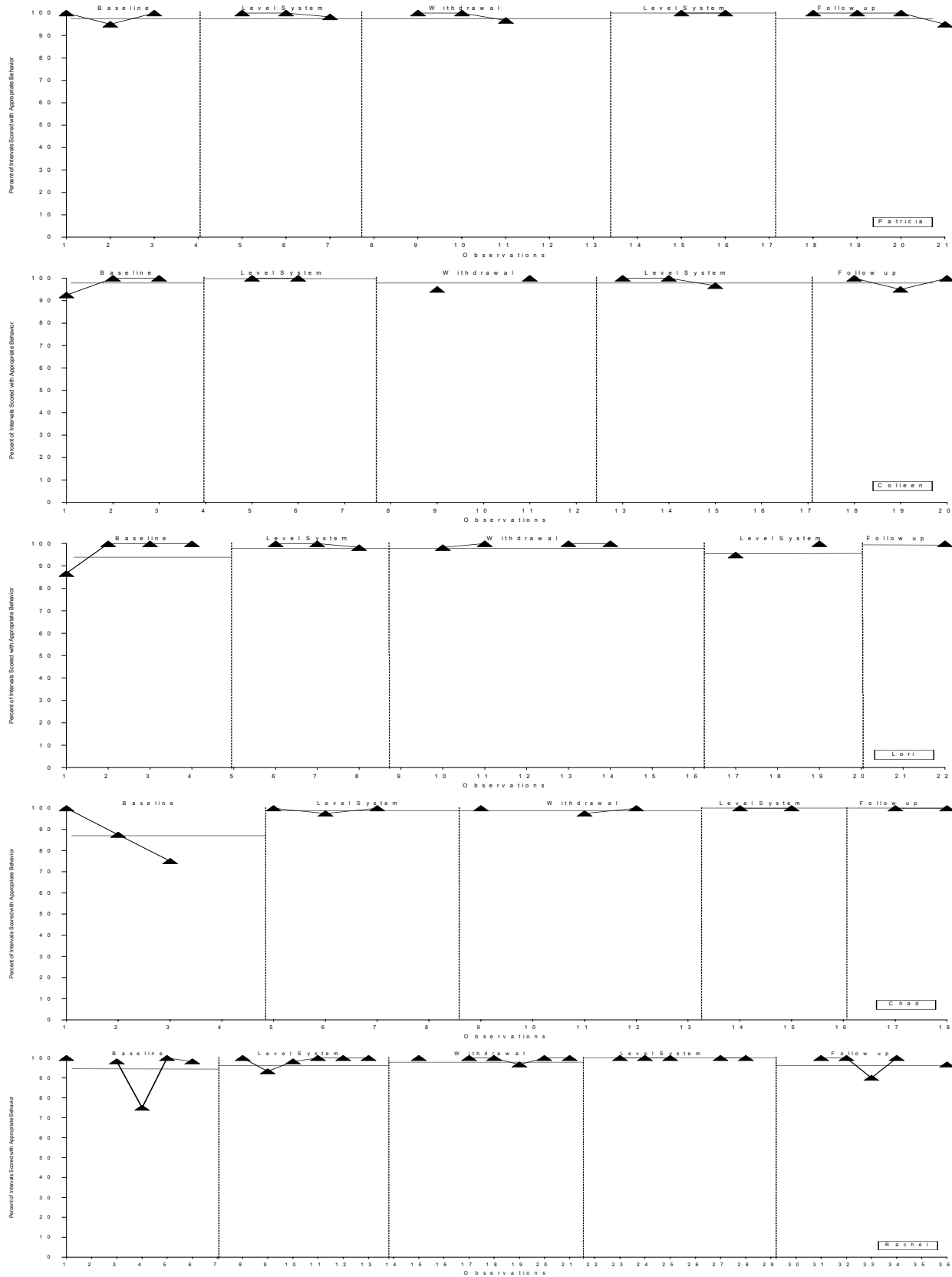


Figure 5. Percentage of intervals scored with appropriate behavior exhibited by John in each teacher's class with horizontal lines indicating means for each phase.

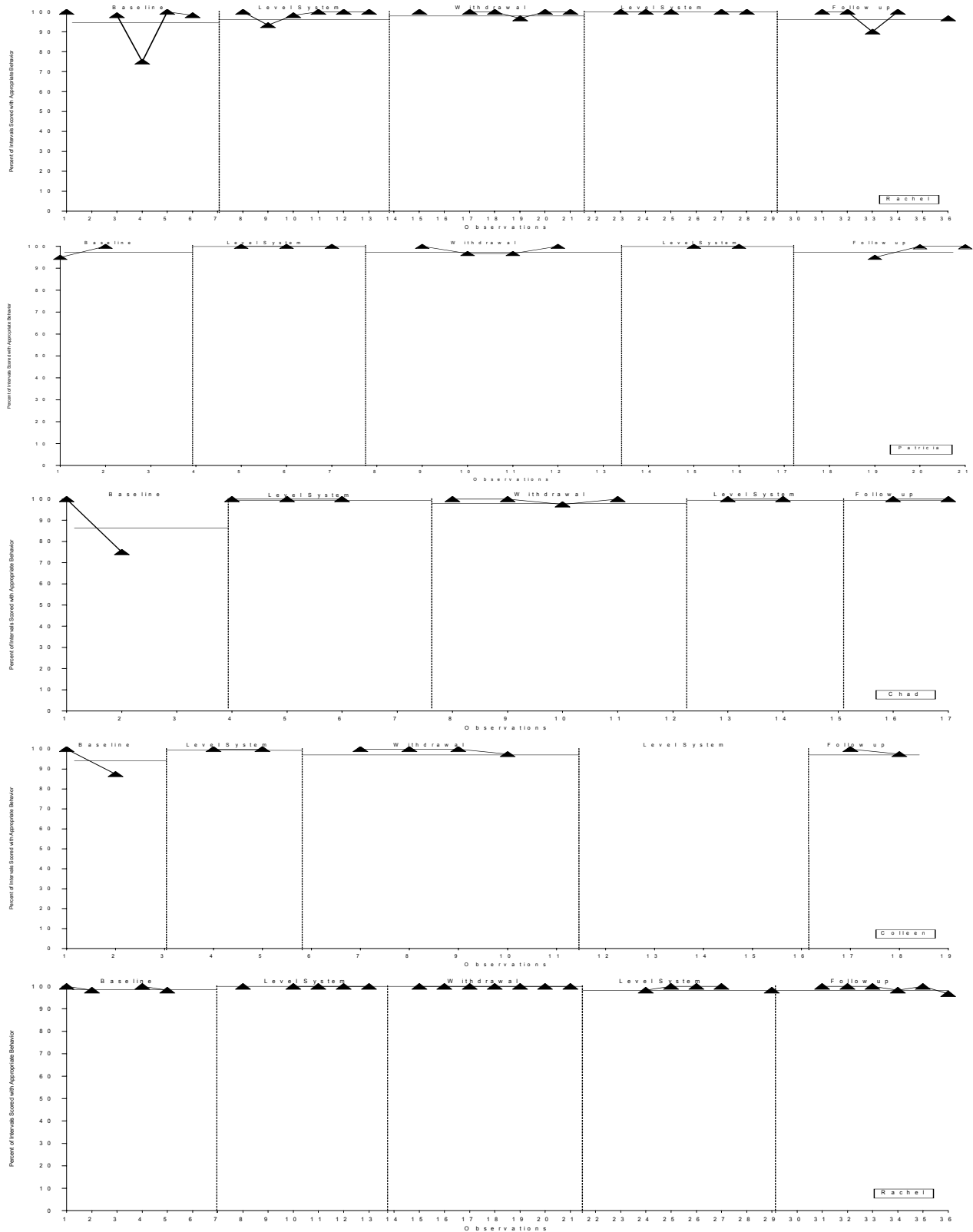


Figure 6. Mean percentage of intervals scored with appropriate behavior exhibited by each participant across teachers.

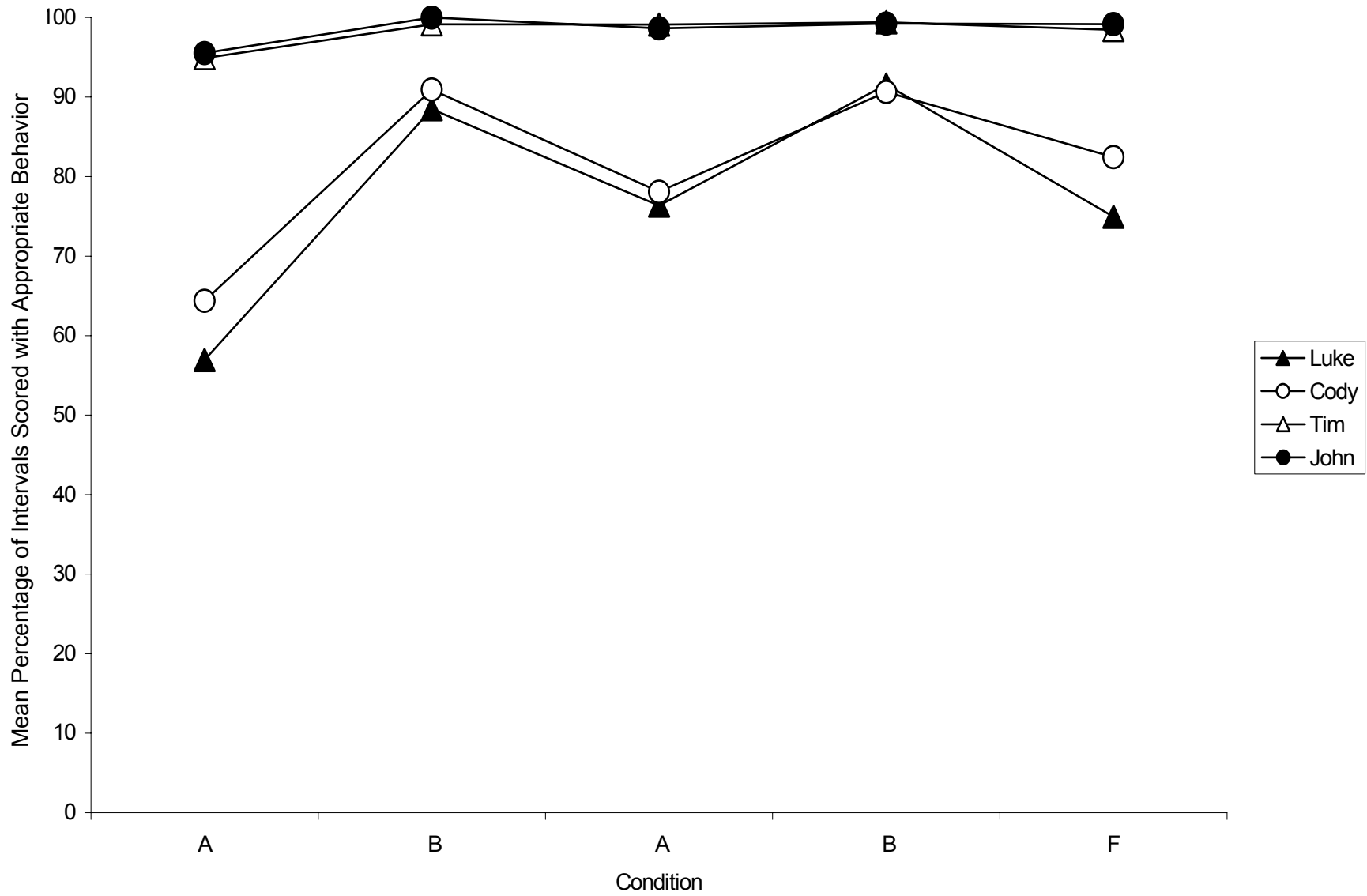


Figure 7. Mean percentage of intervals scored with appropriate behavior exhibited by each participant in each teacher's class.

