Article

Hammill Institute on Disabilities

Topics in Early Childhood Special Education
2017, Vol. 36(4) 230–241

DOI: 10.1177/0271121416649935
tecse.sagepub.com
SAGE

A Review of the Use of Group Contingencies in Preschool Settings

Elizabeth A. Pokorski, MS[1], Erin E. Barton, PhD[1], and Jennifer R. Ledford, PhD[1]

Abstract

Individual contingency management systems have been used successfully to improve behaviors in school settings—including preschools—but often come with associated challenges in time and personnel management. Group contingencies, in the form of independent, interdependent, and dependent contingencies, have been used in preschools to address these challenges. Although current reviews exist on the positive effects of these interventions with older students, none exist to date on the effects on preschool children. The purpose of this review was to synthesize the research on the use of group contingencies in preschool classrooms. Clear trends emerged across studies regarding characteristics of participants, type of contingency used, intervention components, and rewards.

Keywords

preschool, behavioral interventions, group contingencies, contingency management, classroom management

Contingency management systems have been used to successfully improve behaviors since the principles of applied behavior analysis (ABA) were first introduced in mid-20th century and have been shown to be effective in school settings with diverse populations (Cooper, Heron, & Heward, 2007). These systems create targeted, meaningful stimulus-response contingencies and might be helpful for children for whom typically available classroom reinforcers are not effective. Individual contingencies can be effective for modifying multiple behavioral subsets (e.g., inappropriate/appropriate behavior, stuttering, social skills) when used in preschool settings, either informally or within more structured individualized behavior plans (Bryant, Herndon Vizzard, Willoughby, & Kupersmidt, 1999; Chandler, Lubeck, & Fowler, 1992; Onslow, Andrews, & Lincoln, 1994). Although these individual contingency systems might be effective in producing a desired behavioral change, they often come with substantial challenges (Litow & Pomroy, 1975). Individual contingencies typically require intensive individualized attention, including consistent individual reinforcement and delivery of rewards (Albers & Greer, 1991). These components require considerable teacher attention to individual children, which might result in diverted attention from other children in the class. In addition, individual contingency systems might not be viewed as socially acceptable classwide interventions because they differentiate specific children, potentially negatively, from their peers. Group contingencies have the potential to alleviate some of these challenges while providing benefits to both targeted and non-targeted children (Cooper et al., 2007).

Group Contingencies

Group contingencies are those in which a shared consequence applies to all members of a group and are based on the performance of one, some, or all members of that group. There are three types of group contingencies: independent, interdependent, and dependent (Litow & Pomroy, 1975). In the case of *independent group contingencies*, all members of a group are held to the same criterion, but each member's performance toward that criterion is measured individually. Consequently, only those group members who individually reach criterion receive reinforcement. Group members share the same criterion in *interdependent group contingencies* as well, but all members of the group must meet the standard for any of the members to receive reinforcement. Thus, even if only one member does not meet criterion, none of the members receive reinforcement. Finally, *dependent group contingencies* measure the performance of select members of the group only: If these members meet criterion, all members of the group receive reinforcement. Each

[1]Vanderbilt University, Nashville, TN, USA

Corresponding Author:
Elizabeth A. Pokorski, Vanderbilt University, 230 Appleton Place, Peabody 228, Nashville, TN 37203, USA.
Email: elizabeth.a.pokorski@vanderbilt.edu

of these contingency types has its individual strengths and limitations (Poplin & Skinner, 2003).

Group contingencies are regularly used in elementary and middle school settings and are often a component of positive behavioral interventions and supports (PBIS) systems (Molloy, Moore, Trail, Van Epps, & Hopfer, 2013). They have been shown to be effective with diverse age groups and with diverse populations—including students with various disabilities (Cooper et al., 2007). Independent, interdependent, and dependent group contingencies have been shown to be effective, but substantially more research exists on the positive effects of interdependent contingencies (Litow & Pomroy, 1975; Little, Akin-Little, & O'Neill, 2015; Maggin, Johnson, Chafouleas, Ruberto, & Begrren, 2012). Group contingencies may take many forms—including simple verbal contingencies, the good behavior game, level charts, and token economies—and the criterion and reinforcement (types and schedules) can be modified to meet changing child and classroom needs. Furthermore, individual contingencies can be added to pre-existing classwide group contingencies for non-responders or children who might need more immediate or frequent reinforcement (Little et al., 2015).

PBIS in Early Childhood Programs

Group contingencies have the potential to be an important component of PBIS in early childhood programs. In recent years, there has been an increased awareness of the importance of addressing social competence and challenging behaviors during preschool, before children start kindergarten (Dunlap et al., 2006; Hemmeter, Ostrosky, & Fox, 2006). Preschool teachers reported that the most demanding aspect of their job was addressing challenging behavior (Lambert, O'Donnell, Kusherman, & McCarthy, 2006) and were unlikely to use effective strategies to address or prevent the challenging behaviors (Hoover, Kubicek, Rosenberg, Zundel, & Rosenberg, 2012; Vinh, Strain, Davidon, & Smith, 2016). Furthermore, Gilliam (2005) found that preschool children who demonstrated challenging behavior were expelled from their preschool programs at alarming rates. Providing teachers with appropriate and feasible interventions to address challenging behavior in the classroom, such as group contingencies, might decrease challenging behaviors, increase social competence, and reduce preschool expulsions (Hemmeter, Fox, Jack, & Broyles, 2007).

Acknowledgment systems, or targeted positive reinforcement for appropriate behavior, are a core component of PBIS systems in schools (Lewis, Powers, Kelk, & Newcomer, 2002; Molloy et al., 2013). Although the principles are the same, in early childhood programs acknowledgment systems—such as group contingencies—need to be developmentally appropriate, specifically taught, and often paired with descriptive and enthusiastic feedback and immediate reinforcement for individual children (Dunlap & Fox, 2015). The procedural flexibility of group contingencies allows early childhood teachers to adapt and modify these systems to meet the specific and often heterogeneous needs of multiple children in one classroom (Maggin et al., 2012). Furthermore, group contingencies can address multiple behavioral subsets, including social skills *and* challenging behaviors. Developmentally appropriate group contingencies have the potential to lead to greater engagement and learning for all children in preschool classrooms. Group contingencies might be used within early childhood program-wide PBIS systems to reduce challenging behaviors, promote appropriate behavior, or target social skill development.

Current Research

Although meta-analyses of group contingencies have been conducted (Little et al., 2015; Maggin et al., 2012), no such analyses have been done specifically on the use of group contingencies with preschool children. Maggin et al. (2012), for example, concluded that although there were 74 individual and 36 group cases with strong to moderate evidence supporting their use, the research on group contingencies has primarily been conducted with White, male students in elementary school with average to low academic achievement. Little et al. (2015) reported strong overall and individual effect sizes for group contingencies using the 182 single case studies identified for their review. However, their results should be interpreted with considerable caution given they calculated effect sizes using analysis methods that remain significantly limited and violate multiple assumptions when applied to single case research (Shadish, 2014; Shadish, Hedges, Horner, & Odom, 2015; Wolery, 2013). Similarly, although reviews of interventions for preschool children exhibiting challenging behavior were available (Bryant et al., 1999; Conroy, Dunlap, Clarke, & Alter, 2005; Dunlap & Fox, 2011), no reviews focused on the use of group contingencies to modify multiple behavioral subsets.

This review synthesized the research on the use of group contingencies in preschool classrooms. Such a review is needed to evaluate the efficacy of group behavior management systems for young children, whose distinct developmental differences might produce differential outcomes from older children. Given the current focus on PBIS for young children (Dunlap & Fox, 2015; U.S. Department of Health and Human Services and U.S. Department of Education, 2015), a review of existing studies identifying critical components of group contingency systems is essential and might lead to specialized systems in the future. For example, the U.S. Department of Health and Human Services and U.S. Department of Education (2015) recently

issued a policy statement on expulsion and suspension in early childhood settings. They recommended using preventive practices to teach appropriate behaviors and facilitate social emotional competence in young children and highlighted early childhood PBIS as a promising approach. Group contingencies might be an effective, feasible component of program-wide PBIS systems.

Specifically, this review addressed the following questions in relation to preschool classrooms:

Research Question 1: With whom have group contingencies been used?
Research Question 2: Which target behaviors have been identified when group contingencies were used?
Research Question 3: Are differential outcomes related to variations in contingencies or rewards?
Research Question 4: Do studies demonstrate sufficient methodological rigor (i.e., meet design standards and quality indicators) to permit evaluation of whether group contingencies should be a recommended practice in early childhood settings?

Method

Search Procedures

Literature searches were conducted through the PsycINFO and ERIC databases using the terms *group contingency, contingency management*, and *token economy programs*, paired with the terms *preschool, head start*, and *early childhood*. This search provided 233 hits, which were abstract-screened for inclusion based on the following conditions: (a) a peer-reviewed article or dissertation, (b) the use of a group contingency as the primary intervention for improving the dependent variable, and (c) the inclusion of preschool-aged participants (aged 2–5 years). This screening resulted in 16 potentially eligible sources. Ancestral searches were conducted on these sources and pertinent reviews located through the original search, resulting in the identification of one additional study. A final full text screening was conducted to exclude any sources that met the following conditions: (1) sources that were not experimental (i.e., included a minimum of three opportunities to demonstrate behavior change at three different points in time for single case research studies), (2) sources that compared two types of group contingency interventions rather than an intervention versus control, or (3) sources that were dissertations later published as peer-reviewed articles. Five articles were excluded because they were non-experimental, two articles were excluded because they compared two types of group contingencies, and one article (a dissertation) was excluded because it was later published (Murphy, Theodore, Alioso, Alric-Edwards, & Hughes, 2007, was identified for this review). An attempt was made to contact the authors of the included dissertations to verify their publication status, but none were located.

Nine sources met final inclusion criteria for this review. One source was considered two studies because each used a different type of group contingency with a separate group of children (Herman & Tramontana, 1971). These two studies are referred to individually as Study 1 and Study 2. In sum, 10 studies were identified and analyzed.

Coding Procedures

Descriptive characteristics of identified studies were coded using a systematic coding protocol. The first author coded all entries, and a graduate student in special education coded 30% of the studies ($n = 3$) to calculate interobserver agreement (IOA). For each study, information was extracted for 28 variables comprising the following areas: study description, dependent variables, intervention characteristics and components, results, and study rigor. IOA was 90.9% across coded variables (Agreements / [Agreements + Disagreements] × 100), with a range of 90.9% to 97% across studies. All coding discrepancies were discussed, and final decisions were made through consensus.

To provide information on for whom and for which behaviors group contingencies have been used, each study was coded for the following variables: research design, number of participants, participant and setting characteristics (age, disability status, presence of behavioral challenges, classroom type, and activity type), and dependent variable topographies and measurement. To analyze the efficacy of the treatment, each study was coded for the following variables: type of group contingency, primary dependent variable and any additional variables measured, unit of analysis (child or group), and teacher behaviors reported present during baseline: praise, reprimands, redirections, time-outs, or other. Each study was further coded to determine which group contingency intervention components were present, including pre-intervention participant training, a visual component (token board, level chart, checklist), reward type and category (social, tangible, activity, edible), and reward selection method.

The rigor and results of each study were separately analyzed. First, methodological features were coded to assess study quality and internal and external validity. This included examining the measurement of IOA, procedural fidelity, generalization, and social validity. Second, a systematic protocol based on Kratochwill et al.'s (2013) single case design standards and Horner and colleagues' (2005) single case quality indicators was used to determine if studies adhered to acceptable contemporary single case methodology. Third, a systematic visual analysis was conducted to identify the strength of the relation between the intervention and outcome variable(s). The systematic visual analysis examined the following data characteristics: level, trend,

Table 1. Study Characteristics.

Study	Design	*n*	Participants				Setting activity	
			Age[a]	Typical development	ASD	Challenging behavior[b]		
Filcheck (2004) dissertation	A-B-A-B	4	4	•		•	General	L
Herman and Tramontana (1971) Study 1	A-B-A-B	3	5–6	•			General	N
Herman and Tramontana (1971) Study 2	A-B-A-B	3	5–6	•			General	N
Hunt (2013) dissertation	MB[c]	3	4–5	•			General	L
Kohler et al. (1995)	A-B-A-B	9	4	•	•		General	P
Ling and Barnett (2013)	A-B-A-B	7	PK	•			General	L
Maus (2007) dissertation	A-B-A-B	7	3–4		•	•	Self-Contained	L
Murphy, Theodore, Alioso, Alric-Edwards, and Hughes (2007)	A-B-A-B	8	3–5	•			General	L
Reitman, Murphy, Hupp, and O'Callaghan (2004)	A-B-A-B	3	PK	•		•	General	L
Swiezy, Matson, and Box (1992)	MB[d]	4	4–5	•			General	G, P

Note. ASD = autism spectrum disorder; L = large group instruction; G = small group game; N = nap; MB = multiple baseline; P = play; • = variable present.
[a]PK = prekindergarten, no ages reported. [b]Challenging behaviors were evaluated and quantified using standardized assessments. [c]MB across classrooms. [d]MB across participant pairs and therapists.

variability, immediacy of effect, and overlap of data within and across conditions and tiers (Gast & Spriggs, 2014). The first and second authors independently coded all figures for all dependent variables, resulting in IOA of 88% (Agreements / [Agreements + Disagreements] × 100). All coding discrepancies were discussed and resolved through further visual analysis, and final decisions were made through consensus. Fourth, an adapted version of Reichow's (2010) success estimate ratio for single case research (i.e., successful demonstrations divided by total attempted demonstrations) was applied to each study to identify the total number of demonstrations of a functional relation. His success estimate was adapted so demonstrations were conceptualized as demonstrations of a functional relation for one dependent variable (with at least three opportunities to document behavioral change at three different points in time), rather than demonstrations of behavioral change (i.e., change from one condition to the adjacent condition). In this manner, 1 of 1 meant that in an A-B-A-B study, each condition change resulted in behavior shifting in the hypothesized direction for one dependent variable. Conversely, an A-B-A-B study where the behavior improved with intervention, but did not reverse when the intervention was withdrawn, would receive a 0 of 1. This adaption allowed for an efficient analysis of the presence or absence of functional relations.

Results

The 10 sources included in this review were published between 1971 and 2013 and consisted of six peer-reviewed publications and three dissertations; none of the articles included any shared authors. All studies occurred within preschool programs during a range of activities (e.g., large group instruction, play, nap; see Table 1).

Participants and Settings

Together, the studies included 51 children (3–9 participants per study) with an age range of 3 to 6 years. Two studies reported only that children were "preschool aged" (Ling & Barnett, 2013; Reitman, Murphy, Hupp, & O'Callaghan, 2004), which precluded calculating a mean age of children included in this review. The majority of studies did not report the developmental status of participants, but two studies reported inclusion of participants diagnosed with autism (Kohler et al., 1995; Maus, 2007). Nine studies reported that their participants displayed challenging behaviors in the classroom, but only three used standardized assessments—either *Conners' Teacher Rating Scale-Revised: Long* (1997) or the *Vineland Adaptive Behavior Scales* (Sparrow, Balia, & Cicchetti, 1985)—to document and quantify these behaviors (Filcheck, 2004; Maus, 2007; Reitman et al., 2004). A total of 20 teachers and instructional assistants across eight studies served as the implementers; in Herman and Tramontana (1971; Studies 1 and 2), researchers (the authors) implemented the contingency systems.

Nine of 10 studies (i.e., excluding Maus, 2007) were conducted in a regular preschool setting. Class sizes ranged from seven to 20, with the majority of the classes containing between 17 and 20 children. Group contingencies were most often implemented in large groups during instruction ($n = 6$), and also were implemented at naptime (Herman & Tramontana, 1971; Studies 1 and 2), during free play contexts (Kohler et al., 1995; Swiezy, Matson, & Box, 1992), and embedded within small group games (Swiezy et al., 1992).

Table 2. Intervention Characteristics.

Study	Contingency type	DV					
		Primary DV	Measurement (primary DV)	Other DV	Operationally defined	Unit of analysis	Baseline
Filcheck (2004)	Independent	AB	PIR − 15 s		•	Child	R, RD, TO
Herman and Tramontana (1971) Study 1	Independent	IB	PIR − 10 s		•	Group	NR
Herman and Tramontana (1971) Study 2	Interdependent	IB	PIR − 10 s		•	Group	NR
Hunt (2013)	Interdependent	IB	PIR − 15 s	Engagement	•	Group, Child	NR
Kohler et al. (1995)	Interdependent	SS	PIR − 6 s	Teacher behavior	•	Child	O
Ling and Barnett (2013)	Interdependent	IB	PIR − 15 s	Teacher behavior, Engagement		Group	R, RD, TO
Maus (2007)	Independent	IB	PIR − 15 s		•	Child	NR
Murphy, Theodore, Alioso, Alric-Edwards, and Hughes (2007)	Interdependent	IB	PIR − 15 s		•	Child	R, TO
Reitman, Murphy, Hupp, and O'Callaghan (2004)	Dependent	IB	PIR − 10 s		•	Child	NR
Swiezy, Matson, and Box (1992)	Interdependent	AB	% Compliance		•	Group	NR

Note. DV = dependent variables. AB = appropriate behavior; PIR = partial interval recording system; R = reprimands; RD = redirections; TO = time-outs; IB = inappropriate behavior; NR = not reported; SS = social skills; O = other; • = variable present.

Dependent Variables

In the majority of studies, researchers examined the effect of the intervention on appropriate (n = 2) or inappropriate behavior (n = 7). In other studies, researchers measured the social interactions (Kohler et al., 1995) or engagement of participants (Hunt, 2013; Ling & Barnett, 2013). In addition, teacher behavior was measured in two studies (Kohler et al., 1995; Ling & Barnett, 2013). All researchers operationally defined dependent variables in included studies, with the exception of Ling and Barnett (2013). Six of 10 studies used individual participant data as the unit of analysis, and five used group data (either small group or whole-class). One study reported individual results for target children and for their classrooms as a whole (Hunt, 2013). Nine of the 10 studies used a PIR system to measure the primary dependent variable. Swiezy et al. (1992) measured percentage of compliance with direct instructions (Number of Compliance Behaviors / Number of Direct Instructions × 100).

Study Designs

All included studies utilized single case design methodology. The most frequently used research design was an A-B-A-B withdrawal design. Two studies used a multiple baseline design (i.e., across classrooms, Hunt, 2013; across participant pairs and therapists, Swiezy et al., 1992). Reitman et al. (2004) classified their study as an alternating treatments with reversal design; however, because the two alternating treatments were iterations of the same intervention (i.e., in one treatment, the target participants were the dependent members of the group intervention and in the other, they were nondependent members), it was designated an A-B-A-B withdrawal design for the purpose of this review.

Studies typically reported that teachers conducted "business as usual" during the baseline condition. Further details, including the teacher's reinforcement and/or discipline procedures, were only provided for four of the studies; the remaining six reported no information on the teacher's behavior during baseline. The most common classroom management strategies reported during baseline were reprimands and time-outs (n = 3) followed by redirections (n = 2). Kohler et al. (1995) reported that the teacher "did not interact with the children except to resolve conflicts over play materials or roles" (p. 19). The use of praise or other positive reinforcement was not reported in any study.

Intervention Characteristics and Components

The most common type of group contingency used was interdependent (six of 10 studies) followed by independent group contingencies (3 of 10 studies). Reitman et al. (2004) used a dependent contingency, which was the only example of a dependent contingency within this review (see Table 2).

Contingency. Although researchers across all studies included teacher training prior to intervention, only four trained children in the use of the group contingency prior to intervention (see Table 3). Child trainings included the behaviors required to meet the group contingency criterion during the intervention—the classroom rules for all studies but one (Kohler et al., 1995, who taught social skills)—and on the procedure for receiving tokens and rewards. The remaining six studies presented the rules and contingency criterion to the participants at the start of

Table 3. Intervention Components.

Study	Contingency		Rewards		
	Training	Visual component	Type	Category[a]	Selection
Filcheck (2004)		•	Mystery	S, T, A, E	Teacher
Herman and Tramontana (1971) Study 1	•	•	NR	T	NR
Herman and Tramontana (1971) Study 2	•	•	NR	T	NR
Hunt (2013)		•	Known	T	Child
Kohler et al. (1995)	•	•	NR	T	NR
Ling and Barnett (2013)			Mystery	T, A	NR
Maus (2007)			Mystery	T, A, E	Teacher
Murphy, Theodore, Alioso, Alric-Edwards, and Hughes (2007)			Mystery	T, A	Teacher
Reitman, Murphy, Hupp, and O'Callaghan (2004)	•	•	Mystery	S, T, A	NR
Swiezy, Matson, and Box (1992)		•	Known	S, E	Teacher

Note. S = social; T = tangibles; A = activities; E = edibles; NR = not reported; • = variable present.

each session. The studies were divided between those that included a displayed visual component for child participants (to track appropriate behaviors) and those that did not. Studies including a visual component used a level system (Filcheck, 2004), token board (Kohler et al., 1995; Swiezy et al., 1992), rewards target game (Reitman et al., 2004), checklist on a whiteboard (Hunt, 2013), and empty bin to be filled with ping-pong balls (Herman & Tramontana, 1971, Studies 1 and 2). Those that did not include a visual component used a checklist or counter, and an adult tracked inappropriate behaviors without sharing results with children.

Rewards. Seven of the studies stated whether the rewards were known to the participants prior to each session ($n = 2$) or were unknown "mystery" rewards ($n = 5$). No information on reward type was provided for the remaining three studies, as seen in Table 3. The reward categories used included social, tangible, activity, and edible. All but one study (Swiezy et al., 1992) included tangibles as rewards, which consisted of small items such as stickers, stamps, inexpensive toys, and books. Six used activities as rewards (e.g., parachute play, outside time, classroom games, dance party), three provided edibles as rewards, and three studies included a social reward (praise) given directly after an appropriate behavior was demonstrated (Filcheck, 2004; Reitman et al., 2004; Swiezy et al., 1992). Many of the studies did not report any information on how the rewards were selected. However, four did state or indicate teacher selection and one reported that the participants themselves choose an item from the "treasure box" (Hunt, 2013). No study reported the use of preference or reinforcer assessments for reward selection.

Study Outcomes

Data from each study were visually analyzed, and an adapted version of Reichow's success estimate ratio for single case research (i.e., successful demonstrations divided by attempted demonstrations) was applied to each study. Studies varied widely in the number of attempted demonstrations of a functional relation, from one (Herman & Tramontana, 1971, Studies 1 and 2; Swiezy et al., 1992) to eight (Murphy et al., 2007). Across 10 studies, nine of 34 attempted demonstrations of a functional relation (26%) were successful. Five studies demonstrated a functional relation for 100% of attempted demonstrations (ranging from 1 of 1 to 3 of 3), and one study demonstrated a functional relation for one of three attempted demonstrations (Reitman et al., 2004). Table 4 provides a visual representation of the number of attempted demonstrations of a functional relation (for the primary dependent variable) per study and whether each resulted in a successful demonstration.

Five of six studies that measured challenging behavior without quantifying participants' behavioral challenges resulted in a demonstration of a functional relation. One of three studies that measured challenging behavior and quantified participants' behavioral challenges resulted in a demonstration of a functional relation (Reitman et al., 2004). In addition, six of eight studies that included typically developing children resulted in 100% successful demonstrations of a functional relation, whereas one of two studies including children with autism did so (Kohler et al., 1995). Of the three studies that measured engagement or teacher behavior as a secondary dependent variable, none resulted in a demonstration of a functional relation. Six of seven studies that included a visual component demonstrated a functional relation (Herman & Tramontana, 1971, Studies 1 and 2; Hunt, 2013; Kohler et al., 1995; Reitman et al., 2004; Swiezy et al., 1992) compared with zero of three studies that did not include a visual component (Ling & Barnett, 2013; Maus, 2007; Murphy et al., 2007).

Table 4. Demonstration of Functional Relation for Primary DV by Attempt.

Study	Demonstration Attempt								Successful/attempted
	1	2	3	4	5	6	7	8	
Filcheck (2004)	—	—	—	—					0/4
Herman and Tramontana (1971) Study 1	•								1/1
Herman and Tramontana (1971) Study 2	•								1/1
Hunt (2013)	•	•							2/2
Kohler et al. (1995)	•	•	•						3/3
Ling and Barnett (2013)	—	—	—	—					0/4
Maus (2007)	—	—	—	—	—	—	—		0/7
Murphy, Theodore, Alioso, Alric-Edwards, and Hughes (2007)	—	—	—	—	—	—	—	—	0/8
Reitman, Murphy, Hupp, and O'Callaghan (2004)[i]	—	—	•						1/3
Swiezy, Matson, and Box (1992)	•								1/1

Note. DV = dependent variables; • = Demonstration of functional relation; — = No demonstration of functional relation.

Table 5. External Validity.

Study	IOA		Procedural fidelity		Generalization	Maintenance	Teacher acceptability
	≥20% sessions	≥80% agree	≥20% sessions	≥80% agree			
Filcheck (2004)	•	•	•	•	•	•	A
Herman and Tramontana (1971) Study 1	•	•			•		NR
Herman and Tramontana (1971) Study 2	•	•			•		NR
Hunt (2013)	•	•	•	•			A
Kohler et al. (1995)	•						NR
Ling and Barnett (2013)	•	•	•	•			A
Maus (2007)	•	•					A
Murphy, Theodore, Alioso, Alric-Edwards, and Hughes (2007)	•	•					A
Reitman, Murphy, Hupp, and O'Callaghan (2004)	•	•					V
Swiezy, Matson, and Box (1992)	•	•			•		NR

Note. IOA = interobserver agreement; A = acceptable; NR = not reported; V = variable acceptability; • = variable present.

Study Rigor

Studies generally displayed low to moderate quality, as measured by the presence and quality of IOA and procedural fidelity data, generalization and maintenance assessment, and social acceptability (Horner et al., 2005). All studies reported collecting IOA data for 20% or more of sessions and participants, and all but Kohler et al. (1995) demonstrated agreement at or above 80% (see Table 5). However, only Hunt (2013) and Ling and Barnett (2013) specifically reported that IOA data were collected reliably across all relevant conditions and participants.

In six studies, researchers collected data on perceived teacher acceptability using a version of the Intervention Rating Profile developed by Witt and Martens (1983), and nearly all found the intervention to be very acceptable (n = 5). Only Reitman et al. (2004) reported variable acceptability, noting that the teacher in their study found the intervention highly acceptable for one participant, moderately so for the second, and not acceptable for the third. Although teachers generally reported the intervention to be acceptable, teachers who reported high acceptability in the one study in which authors attempted to collect maintenance data did not elect to continue the intervention (Filcheck, 2004).

In general, the studies did not meet contemporary recommendations for procedural fidelity as described by Horner and colleagues (2005); only Filcheck (2004), Hunt (2013), and Ling and Barnett (2013) collected fidelity data for 20% or more of sessions and reported 80% or greater fidelity. Maus (2007) relied on a teacher self-report to track procedural adherence, but these data were not included due to their potential bias. No author stated that fidelity measures were collected across all relevant conditions and participants.

Researchers in four studies measured generalization. For example, researchers found effects generalized across teachers (Filcheck, 2004; Swiezy et al., 1992). However, group contingencies did not generalize across settings (Herman & Tramontana, 1971, Studies 1 and 2).

Discussion

This review was conducted to evaluate the literature on the use of group contingencies in preschool settings to determine with whom and for which behaviors group contingencies have been used, under what conditions group contingencies have been shown to be effective, and if studies demonstrate adequate methodological rigor. The authors analyzed 28 components across five areas to characterize study features, outcomes, rigor, and credibility of results. Patterns were seen across studies in the types of group contingencies used, qualities of rewards, with whom contingencies were used, and in what settings contingencies were used. Studies generally did not report procedures during "business as usual" baseline conditions, did not report procedural fidelity data, and did not document and quantify behavioral challenges adequately. However, other quality indicators were present and acceptable (i.e., IOA, participant and intervention descriptions, research design), indicating moderate methodological rigor across studies and permitting an analysis of their efficacy.

Overall Analysis

Group contingencies have the potential to be effective in decreasing inappropriate behavior or increasing appropriate behavior with typically developing children both with and without documented behavioral challenges, particularly when visuals are used. Positive behavioral changes were observed in six studies, and no negative effects were reported in any studies. Overall, these findings supported other recent reviews of group contingencies (Little et al., 2015; Maggin et al., 2012). In addition, in this and other reviews, group contingencies showed positive behavior change of classes as a whole. However, because the data across these studies were generally based on the presence or absence of *any* child in the class exhibiting a behavior of interest, it is not apparent if effects were due to a change in behavior across the majority of children or if they were due to a substantial change in the behavior of one or more target children who displayed high occurrences of problem behavior. It is interesting to note that studies that documented and measured behavioral challenges in children tended to result in poorer outcomes than those that did not (Filcheck, 2004; Maus, 2007; Reitman et al., 2004). This might indicate that for children with more substantial behavioral challenges, group contingencies might not offer the same benefits as individual contingencies. Adding an individual contingency to the classwide group contingency or selecting rewards based on a preference or reinforcer assessment of specific children might produce more positive outcomes.

In the current review, evidence suggested that group contingencies were effective for some, but not all, children with autism and most children with typical development. Effects on the social skills of children also were positive; however, only one study assessed this variable in the context of a group contingency (Kohler et al., 1995). The addition of a group contingency resulted in improvements in appropriate/inappropriate behavior across multiple studies, which suggests positive changes in social skills.

An effect was not indicated across all attempted demonstrations in all studies. However, in some cases, this was due to data failing to return the level of behavior seen in the baseline condition when the intervention was withdrawn (Filcheck, 2004, Child 1 and 2; Maus, 2007, Child 3, 5, and 7; Murphy et al., 2007, Child 1, 2, 4, 5, and 7). Although this precludes identification of functional relations, it suggests that the group contingency resulted in positive behavioral changes, even when used briefly, and maintained when the intervention was discontinued. This capacity to modify behavior quickly and result in sustained positive outcomes speaks to the potential social validity of group contingencies. Several specific and important outcomes from this review are discussed in the next section.

Group contingency types. The majority of studies assessed interdependent group contingencies rather than independent or dependent contingencies, but positive effects were seen across all three types of interventions; these results aligned with previous analyses on group contingencies with older participants (Little et al., 2015; Maggin et al., 2012). Independent and interdependent contingencies can arguably be considered the most reasonable for preschool-aged children because both require each individual to reach the predetermined criterion himself or herself to attain the reward; this direct cause-and-effect contingency has the potential to be more meaningful for young children than does a more complex contingency. It is possible that dependent contingencies are as effective with preschool children (Reitman et al., 2004), but the limited number of studies on this contingency type hinders comparisons. It might be that dependent contingencies were selected less often with preschool children because of their relatively young age. Preschool children might not comprehend a contingency if they are not an active part of its success or failure. Another consideration when selecting a dependent contingency is the issue of justice or fairness that arises when failing to reward those whose behavior was appropriate during the condition, even if the criterion was not met by the dependent members (Gresham & Gresham, 1982).

Reward types. Nearly all the studies that reported reward type used mystery rewards rather than known rewards. The rationale for this decision was provided in four of the studies and was based on the hypothesis that unknown rewards may result in greater behavior change than known rewards (Jenson, Rhode, & Reavis, 1992; Moore, Waguespack, Wickstrom, Witt, & Gaydos, 1994). The predominant use of

mystery rewards appears unexpected when considering previous research indicating the positive value of choice in reinforcer selection (Tiger, Hanley, & Hernandez, 2006); however, the value of choice may not be as meaningful if the rewards provided in the no-choice condition are preferred (Fisher, Thompson, Piazza, Crosland, & Gotjen, 1997) or if the rewards provided in both conditions are equally preferred (Waldron-Soler, Martella, Marchand-Martella, & Ebey, 2000). Although additional research is needed on the efficacy of mystery rewards versus known rewards, the choice to use either did not appear to affect the results: Both conditions produced equally effective results across the included studies. However, because only three of six studies that produced a demonstration of a functional relation reported reward type, conclusions on the impact of reward type on the success of the intervention are limited. In addition, nine studies used tangible rewards, five used activities, and only three identified social (praise) rewards. Researchers generally did not indicate why they chose the reward categories that they did, but because none of the studies reported using preference or reinforcer assessments prior to the intervention, it is likely that reward categories were selected based on perceived child preferences, economic considerations, or feasibility.

Visual components. The presence of a visual component might have positively affected the results of the group contingency intervention. The majority of studies that included a visual component demonstrated a functional relation, whereas no demonstrations of a functional relation were seen in studies that did not include a visual component. Hunt (2013) was the only one of three studies using a checklist to produce successful outcomes, possibly because it was the only study that displayed the checklist for child reference. This indicates it is likely that a group contingency intervention might be less meaningful to preschool-age children when there is not a visual component involved.

Baseline comparisons. A significant shortcoming of the majority of studies was the minimal information reported on the baseline condition. Fewer than half of the studies provided details on the teacher's "business as usual" behaviors, which affected the strength of any conclusions on the intervention's effectiveness. For these studies, it was not clear which aspects of the intervention were consistent with the "business as usual" disciplinary strategies present at baseline and which were distinct changes. Also noted was the contrast between the lack of documented positive reinforcement and the presence of negative disciplinary strategies, such as reprimands and time-outs, during the baseline condition—a result consistent with previous research on the frequency of positive and negative interactions between teachers and students (Jack, Shores, Denny, & Gunter, 1996). In the four studies that did provide baseline descriptions, negative disciplinary behaviors were replaced with consistent, and potentially less aversive, behavioral contingencies; this change in teacher behavior alone may have been enough to produce the positive effects. In addition, three studies noted that praise was provided as an immediate reinforcer during the intervention (in addition to the group contingency). Descriptive praise has been documented to increase a variety of desired behaviors when used consistently in school settings (Jenkins, Floress, & Reinke, 2015), so its addition might have strengthened results.

Measurement procedures. A second major limitation of this body of research was the ubiquitous use of partial interval recording (PIR) systems to measure dependent variables (i.e., challenging behaviors, appropriate behaviors). Although interval systems are often more feasible and reliable than event or duration recording systems when measuring the behaviors of multiple children at one time, recent research suggested that interval systems were likely to be flawed and should not be used (Lane & Ledford, 2014; Ledford, Ayers, Lane, & Lam, 2015; Wood, Hojnoski, Laracy, & Olson, 2016). For example, Ledford et al. (2015) found that interval systems do not provide accurate estimates of duration. In fact, PIR, in particular, was differentially accurate related to the interval size and frequency of behavior. Ledford and colleagues found that for estimating count behaviors using PIR, intervals *longer than* the average duration per occurrence were needed. Interestingly, in this study, five of the nine studies using PIR used 15-s intervals to measure appropriate and inappropriate behavior; three used 10-s intervals, and all measured inappropriate behavior; and one used a 6-s interval to measure social interactions. It is unlikely that the average duration per occurrence of challenging behavior was longer than 10 s or even 15 s. However, none of the studies in the current review reported selecting their interval size based on the frequency or duration of occurrence of the behavior. Furthermore, session lengths varied and were not always reported.

Future Research

The results of this review demonstrate that the current research base is limited. Although results from the included studies indicate the potential for success in decreasing inappropriate behavior with typically developing preschool children through the introduction of a group contingency, the overall methodological quality of these studies affects their validity and precludes definitive conclusions. Limitations were noted throughout the studies in the adequacy of baseline condition descriptions, lack of procedural fidelity information, and inconsistent measurement of generalization and maintenance. However, the scope and quantity of these limitations were not sufficient to prevent analysis of studies and recommendations for practice.

Rigorous research on the use of group contingencies with preschoolers is needed to substantiate the findings of this review. Quality indicators for single case research should guide future research (Horner et al., 2005). For example, although both types have been used extensively, no studies have compared the efficacy of known versus mystery rewards. Future research also might examine conditions under which group contingencies are most likely to be effective. For example, are group contingencies more effective when paired with praise both with and without a secondary reward, or does adding an individual contingency to an existing group contingency increase efficacy for known non-responders. Likewise, the extent to which group contingencies are effective for different preschool populations and different behaviors or children with specific behavioral and functional repertoires should be examined. Given the limited information regarding procedural fidelity, additional research is needed examining effective practices (i.e., coaching) to support high-fidelity teacher implementation of group contingencies. Finally, each of these lines of research should examine the long-term and sustained effects of group contingencies and ideal measurement procedures for accurately measuring primary dependent variables.

Implications for Practice

Although more high-quality research on the use of group contingencies with preschoolers is needed, the implications for practice are quite clear: When used consistently, group contingencies can decrease inappropriate behaviors for typically developing children in preschool settings and should not produce negative effects under these same conditions. Group contingencies might be successfully applied to additional behavioral subsets (e.g., social skills) and populations (e.g., children with disabilities) as well, but the limited research on these uses precludes definitive recommendations. However, because targeting increased positive behavior (rather than decreased negative behavior) and more extensive inclusion across preschoolers is closely aligned with the principles of PBIS, using a group contingency to support these outcomes would be meaningful. Ongoing progress monitoring should be used with group contingencies across all contexts, particularly those outside the scope of this review. Contingencies will likely require regular adjustments in both criterion and rewards when used for a longer period and will be dependent on fluctuating qualities of children and interactions within each classroom. As further exploration of this subject continues, practitioners can feel comfortable applying this intervention with their students, particularly if they use an interdependent group contingency with a visual component and include descriptive praise as an immediate reward for positive behavior.

Authors' Note

Nothing in the article necessarily reflects the positions or policies of the federal government, and no official endorsement by it should be inferred.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research described in this article was supported in part by Grant H325H140001 from the Office of Special Education Programs, U.S. Department of Education.

References

*References marked with an asterisk indicate studies included in this review.

Albers, A. E., & Greer, R. D. (1991). Is the three-term contingency trial a predictor of effective instruction? *Journal of Behavioral Education*, *1*, 337–354.

Bryant, D., Herndon Vizzard, L., Willoughby, M., & Kupersmidt, J. (1999). A review of interventions for preschoolers with aggressive and disruptive behavior. *Early Education and Development*, *10*, 47–68.

Chandler, L. K., Lubeck, R. C., & Fowler, S. A. (1992). Generalization and maintenance of preschool children's social skills: A critical review and analysis. *Journal of Applied Behavior Analysis*, *25*, 415–428.

Conners, C. K. (1997). *Manual for the Conners' Rating Scales-Revised*. Toronto, Ontario, Canada: Multi-Health Systems.

Conroy, M. A., Dunlap, G., Clarke, S., & Alter, P. J. (2005). A descriptive analysis of positive behavioral intervention research with young children with challenging behavior. *Topics in Early Childhood Special Education*, *25*, 157–166.

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

Dunlap, G., & Fox, L. (2011). Function-based interventions for children with challenging behavior. *Journal of Early Intervention*, *33*, 333–343.

Dunlap, G., & Fox, L. (2015). *The pyramid model: PBS in early childhood programs and its relation to school-wide PBS*. Tampa: Pyramid Model Consortium, University of South Florida and Florida Center for Inclusive Communities.

Dunlap, G., Strain, P. S., Fox, L., Carta, J. J., Conroy, M., Smith, B., . . . Sowell, C. (2006). Prevention and intervention with young children's challenging behavior: A summary and perspective regarding current knowledge. *Behavioral Disorders*, *32*, 29–45.

*Filcheck, H. A. (2004). *Evaluation of a whole-class token economy to manage disruptive behavior in preschool classrooms* (Doctoral dissertation). Retrieved from PsycINFO. (Order No. AAI3132942)

Fisher, W. W., Thompson, R. H., Piazza, C. C., Crosland, K., & Gotjen, D. (1997). On the relative reinforcing effects of choice and differential consequences. *Journal of Applied Behavior Analysis*, *30*, 423–438.

Gast, D. L., & Spriggs, A. D. (2014). Visual analysis of graphic data. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed., pp. 176–210). New York, NY: Routledge.

Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems*. Retrieved from http://challengingbehavior.fmhi.usf.edu/explore/policy_docs/prek_expulsion.pdf

Gresham, F. M., & Gresham, G. N. (1982). Interdependent, dependent, and independent group contingencies for controlling disruptive behavior. *The Journal of Special Education*, *16*, 101–110.

Hemmeter, M. L., Fox, L., Jack, S., & Broyles, L. (2007). A program-wide model of positive behavior support in early childhood settings. *Journal of Early Intervention*, *29*, 337–355.

Hemmeter, M. L., Ostrosky, M., & Fox, L. (2006). Social and emotional foundations for early learning: A conceptual model for intervention. *School Psychology Review*, *35*, 583–601.

*Herman, S. H., & Tramontana, J. (1971). Instructions and group versus individual reinforcement in modifying disruptive group behavior. *Journal of Applied Behavior Analysis*, *4*, 113–119.

Hoover, S. D., Kubicek, L. F., Rosenberg, C. R., Zundel, C., & Rosenberg, S. A. (2012). Influence of behavioral concerns and early childhood expulsions on the development of early childhood mental health consultation in Colorado. *Infant Mental Health Journal*, *33*, 246–255.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*, 163–179.

*Hunt, B. M. (2013). *Using the good behavior game to decrease disruptive behavior while increasing academic engagement with a head start population* (Doctoral dissertation). Retrieved from PsycINFO. (Order No. AAI3530731)

Jack, S. L., Shores, R. E., Denny, R. K., & Gunter, P. L. (1996). An analysis of the relationship of teachers' reported use of classroom management strategies on types of classroom interactions. *Journal of Behavioral Education*, *6*, 67–87.

Jenkins, L. N., Floress, M. T., & Reinke, W. (2015). Rates and types of teacher praise: A review and future directions. *Psychology in the Schools*, *52*, 463–476.

Jenson, W. R., Rhode, G., & Reavis, H. K. (1992). *The tough kid book: Practical classroom management strategies*. Longmont, CO: Sopris West.

*Kohler, F. W., Strain, P. S., Hoyson, M., Davis, L., Donina, W. M., & Rapp, N. (1995). Using group-oriented contingency to increase social interactions between children with autism and their peers: A preliminary analysis of corollary supportive behaviors. *Behavior Modification*, *19*, 10–32.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*, 26–38.

Lambert, R. G., O'Donnell, M., Kusherman, J., & McCarthy, C. J. (2006). Teacher stress and classroom structural characteristics in preschool settings. In R. G. Lambert & C. J. McCarthy (Eds.), *Understanding teacher stress in an age of accountability* (pp. 105–120). Charlotte, NC: Information Age Publishing.

Lane, J. D., & Ledford, J. R. (2014). Using interval-based systems to measure behavior in early childhood special education and early intervention. *Topics in Early Childhood Special Education*, *34*, 83–93.

Ledford, J. R., Ayres, K. A., Lane, J. D., & Lam, M. F. (2015). Accuracy of interval-based measurement systems in single case research. *The Journal of Special Education*, *49*, 104–117.

Lewis, T. J., Powers, L. J., Kelk, M. J., & Newcomer, L. L. (2002). Reducing problem behaviors on the playground: An investigation of the application of school wide positive behavior supports. *Psychology in the Schools*, *39*, 181–190.

*Ling, S. M., & Barnett, D. W. (2013). Increasing preschool student engagement during group learning activities using a group contingency. *Topics in Early Childhood Special Education*, *33*, 186–196.

Litow, L., & Pomroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis*, *8*, 341–347.

Little, S. G., Akin-Little, A., & O'Neill, K. (2015). Group contingency interventions with children—1980-2010: A meta-analysis. *Behavior Modification*, *39*, 322–341.

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology*, *50*, 625–654.

*Maus, M. (2007). *Independent group contingencies for reducing disruptive behavior in preschoolers with PDD-NOS* (Doctoral dissertation). Retrieved from PsycINFO.

Miltenberger, R. G. (2001). *Behavior modification: Principles and procedures* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

Molloy, L. E., Moore, J. E., Trail, J., Van Epps, J. J., & Hopfer, S. (2013). Understanding real world implementation quality and "active ingredients" of PBIS. *Prevention Science*, *14*, 593–605.

Moore, L. A., Waguespack, A. M., Wickstrom, K. F., Witt, J. C., & Gaydos, G. R. (1994). Mystery motivator: An effective and time efficient intervention. *School Psychology Review*, *23*, 106–118.

*Murphy, K. A., Theodore, L. A., Aloiso, D., Alric-Edwards, J., & Hughes, T. L. (2007). Interdependent group contingency and mystery motivators to reduce preschool disruptive behavior. *Psychology in the Schools*, *44*, 53–63.

Onslow, M., Andrews, C., & Lincoln, M. (1994). A control/experimental trial of an operant treatment for early stuttering. *Journal of Speech, Language, and Hearing Research*, *37*, 1244–1259.

Poplin, J., & Skinner, C. H. (2003). Enhancing academic performance in a classroom serving students with serious emotional disturbance: Interdependent group contingencies with

randomly selected components. *School Psychology Review*, *32*, 282–295.

Reichow, B. (2010). Development, procedures, and application of the evaluative method for determining evidence-based practices in autism. In B. Reichow, P. Doehring, D. V. Cicchetti, & F. R. Volkmar (Eds.), *Evidence-based practices and treatments for children with autism* (pp. 25–39). New York, NY: Springer.

*Reitman, D., Murphy, M. A., Hupp, S. D. A., & O'Callaghan, P. M. (2004). Behavior change and perceptions of change: Evaluating the effectiveness of a token economy. *Child & Family Behavior Therapy*, *26*, 17–36.

Shadish, W. R. (2014). Analysis and meta-analysis of single-case designs: An introduction. *Journal of School Psychology*, *52*, 109–122.

Shadish, W. R., Hedges, L. V., Horner, R. H., & Odom, S. L. (2015). *The role of between case effect size in conducting, interpreting, and summarizing single-case research*. Washington, DC: Institute for Education Sciences, U.S. Department of Education.

Sparrow, S. S., Balia, D. A., & Cicchetti, D. V. (1985). *Vineland Adaptive Behavior Scales: Classroom edition*. Circle Pines, MN: American Guidance Service.

*Swiezy, N. B., Matson, J. L., & Box, P. (1992). The good behavior game: A token reinforcement system for preschoolers. *Child & Family Behavior Therapy*, *14*, 21–32.

Tiger, J. H., Hanley, G. P., & Hernandez, E. (2006). An evaluation of the value of choice with preschool children. *Journal of Applied Behavior Analysis*, *39*, 1–16.

U.S. Department of Health and Human Services and U.S. Department of Education. (2015, September 14). *Policy statement on the inclusion of children with disabilities in early childhood programs*. Retrieved from: http://www2.ed.gov/about/inits/ed/earlylearning/inclusion/index.html

Vinh, M., Strain, P., Davidson, S., & Smith, B. J. (2016). One state's systems change efforts to reduce child care expulsion: Taking the Pyramid Model to Scale. *Topics in Early Childhood Special Education*. Advanced online publication.

Waldron-Soler, K., Martella, R. C., Marchand-Martella, N., & Ebey, T. L. (2000). Effects of choice of stimuli as reinforcement for task responding in preschoolers with and without developmental disabilities. *Journal of Applied Behavior Analysis*, *33*, 93–96.

Witt, J. C., & Martens, B. K. (1983). Assessing the acceptability of behavioral interventions used in the classrooms. *Psychology in the Schools*, *20*, 510–517.

Wolery, M. (2013). A commentary: Single-case design technical document of the What Works Clearinghouse. *Remedial and Special Education*, *34*, 39–43.

Wood, B. K., Hojnoski, R. L., Laracy, S. D., & Olson, C. L. (2016). Comparison of observational methods and their relation to ratings of engagement in young children. *Topics in Early Childhood Special Education*, *35*, 211–222.