긍정적행동지원연구
Journal of Positive Behavior Support
2024, Vol. 1, No. 2, pp.19-42
https://data.doi.or.kr/10.23288/JisTaP.2024.1.2.2

Teacher-Preferred Group Contingencies with Data-Based Decision Making: Improving Class-Wide Behavior

Fernando Herrera, M.A., BCBA (Medical University of the Americas)
Kwang-Sun Cho Blair, Ph.D., BCBA-D (University of South Florida)*

<요 약>

This study examined the use of teacher-preferred group contingencies, which utilized class-wide data for decision-making to improve classroom behavior in three elementary classrooms. During intervention, the classroom teachers selected an independent or a randomized contingency as their preferred contingency, which they implemented while collecting data on class-wide behavior using a behavior rating scale. The teachers used behavior rating scale data to make decisions about group reinforcement criteria, session length, and types of reinforcers. A concurrent multiple-baseline design across classes was used to evaluate the class-wide behavioral outcomes. Results revealed that the teacher's preferred group contingency, when implemented in conjunction with data-based decision-making, significantly reduced disruptive behavior and moderately increased academic engagement in all three classes. Furthermore, changes in class-wide behavior were maintained at 2-week follow-ups.

<Key words> Group contingency, class-wide, teacher preference, data-based decision-making, disruptive behavior

* 교신저자 : kwangsun@usf.edu

Ⅰ. Introduction

Disruptive behavior in classrooms interrupts instruction and negatively impacts both individual and group learning processes. To address this persistent challenge, teachers are encouraged to employ evidence-based classroom management strategies that can effectively improve student behavior. One such strategy is the use of group contingencies (GCs), which have been shown to improve both academic and behavioral outcomes (Brennan et al., 2024; Maggin et al., 2017). Group contingencies offer strategic alternative to traditional classroom management techniques. By delivering consequences to groups rather than individuals, GCs provide a cost-effective, time-efficient, and practical solution for teachers (Heering & Wilder, 2006; Skinner et al., 2009). This approach significantly reduces the need for teachers to implement individualized interventions for students who repeatedly engage in disruptive behavior (Hulac & Benson, 2010).

Numerous research studies have demonstrated the practical value of GCs in classroom settings. Researchers have consistently reported that teachers find GCs highly acceptable and effective for reducing disruptive behavior and increasing appropriate behavior (Beaver et al., 2023; Bohan et al., 2022; Cihak et al., 2009; Ennis et al., 2016; Donaldson et al., 2018). Moreover, these strategies have shown a strong contextual fit for classroom implementation (Benazzi et al., 2006; McIntosh et al., 2010), making them a promising approach for educators seeking to create more positive and productive learning environments.

The literature identifies four types of GCs: dependent, independent, interdependent, and randomized or unknown contingency (Ennis et al., 2016; Skinner et al., 2002). In dependent GCs, the whole class shares the same expectations, but reinforcement is contingent on the performance of one student or a small group (McKissick et al., 2010). Independent GCs involve applying the same consequences to all students, with reinforcement delivered individually based on each student's behavior (Skinner et al., 2004). Interdependent GCs reinforce the entire class when the group collectively meets a specified behavioral criterion (McKissick et al., 2010), whereas randomized GCs involves reinforcement based on whether the class meets the behavioral criterion of a GC type that is unknown to the class (Kelshaw-Levering et al., 2000).

Research has demonstrated that the four GC types are equally effective in increasing group academic engagement or on-task behavior (Ennis et al., 2016; Hulac & Benson, 2010; Ling et al., 2011) and decreasing problem behavior (Donaldson et al., 2018; Ennis et al., 2016; Hulac &

Benson, 2010). However, students may become unmotivated if they consistently fail to meet the predetermined criteria or if less-preferred reinforcers are selected for the contingency (Skinner et al., 2002). Researchers have suggested that these issues can be minimized when one or more components (e.g., contingency type, criteria, reinforcer) are randomized and selected at the end of the instructional time or by incorporating student preference into reinforcer selection (Alric et al., 2007; Ennis et al., 2016; Little et al., 2010).

The literature has also indicated that in selecting and implementing a GC, classroom teachers should play an active role to enhance and sustain intervention outcomes (Cihak et al., 2009; Heering & Wilder, 2006; Wright & McCurdy, 2012). One way for school-based consultants or researchers to promote teacher involvement in implementing a classroom management strategy is choosing the strategy based on teacher preference. Ennis et al. (2016) reported that, although all four GC types were equally effective in improving classroom behavior to some degree in three elementary classrooms, the implementation of the contingency type preferred by the teachers led to further improvements in class-wide behavior.

With the growing emphasis on regularly collecting and analyzing student monitoring data to inform decision-making in student support (McIntosh et al., 2010), schools have increasingly utilized a variety of data at individual, classroom, and school levels. Data-based decision-making by teachers has been shown to improve instruction and result in better learning outcomes (Black et al., 2004; McNaughton et al., 2012). Despite these benefits, teachers often face challenges in collecting and analyzing monitoring data to guide classroom instruction, posing a major barrier to implementing data-based decision-making in implementing school-wide positive behavioral interventions and supports (PBIS; Ingram et al., 2004; U. S. Department of Education, 2009; Schildkamp & Kuiper, 2010; Wayman, 2005).

The lack of efficient and reliable data collection methods poses a challenge for teachers participating in data-based decision-making to improve student behavior (Chafouleas, 2011). For this reason, researchers have suggested using behavior rating scales that combine the benefits of both a rating scale and systematic direct observation, such as Direct Behavior Rating (Chafouleas et al., 2009) and Individualized Behavior Rating Scale Tool (Ford et al., 2024; Iovannone et al., 2014). Although the length and nature of rating periods may vary, completing the ratings on targeted behaviors immediately following the target academic time period is recommended to facilitate accuracy of ratings and limit bias (Kilgus, 2013).

Therefore, this study focused on facilitating classroom teachers to collect and use data on classroom behavior in a GC intervention to improve classroom outcomes. Specifically, the study examined: (a) the extent to which a teacher-preferred GC that incorporates data-based decision making would reduce class-wide disruptive behavior and increase academic engagement and (b) whether teachers would find implementing the preferred GC with data-based decision making to be acceptable and effective.

Ⅱ. Method

1. Participants and Setting

This study involved three classes, each consisting of one teacher and 18 students from grades 1, 2, and 4, in an urban public elementary school with approximately 790 students. Between 78% and 89% of the students in each class were from minority backgrounds, primarily African American or Hispanic. Between 89% and 94% of the students were receiving free or reduced-price lunch. The classes were recruited through a 2-step recruitment process involving a brief teacher interview and a direct classroom observation. The initial teacher interview aimed to identify potential participation and assess classroom behavioral needs. Classes were included based on the following criteria: (a) the teacher had no prior experience with group contingency and data-based decision making; (d) at least 3 students exhibited disruptive behavior; (e) disruptive behavior occurred daily during at least one instructional time; (f) the teacher implemented weekly academic assessment during the target academic time, and (g) a minimum of 70% of students obtained parental permission and provided assent to participate in the study. Students were excluded from class-wide data collection if their disruptive behavior posed potential danger to themselves or peers, or if they were receiving special education services, requiring tier 3 level individualized behavior interventions.

Teacher interview consisted of questions to identify potential disruptive behavior and the number of students engaging in disruptive behavior during instructional activities. Following the interview, a classroom observation was conducted to verify the number of students exhibiting disruptive behavior and assess the overall class-wide disruption levels. The observation was performed during the 20-60 min academic time identified through the interview as having a high frequency of disruptive

behavior. Data collection utilized a 15-s partial interval recording system to document student disruptive behavior and track the number of students engaging in the disruptive behavior. Across all classes, interviews and observations consistently revealed 3-4 students frequently engaging in disruptive behavior, with overall class-wide disruption levels ranging between 50% – 60% of recorded intervals.

Class 1 was a 1st-grade class with a 30-year-old White female teacher who had a Bachelor of Science degree in Elementary Education and 8 years of teaching experience. The class targeted daily reading workstations, during which students participated in small-group instruction on reading skills and independent work at different stations around the classroom. Class 2 was a 2nd-grade class with a 31-year-old White female teacher with 2 years of teaching experience and a Bachelor of Science degree in Sociology. The class targeted reading, during which students participated in whole group instruction, independent seat work or independent reading on the carpet, and occasional hands-on activities (e.g., making posters). During this period, eight students were receiving additional individual instructional support from the English Language Learners support staff. Class 3 was a 4th-grade class with a 38-year-old female teacher with 10 years of teaching experience. The teacher had a Bachelor of Arts degree in Theatre with a minor in English and a master's degree in Reading Education. The class targeted math. During this time, students participated in whole group instruction and independent seatwork or small group work, and all students received additional academic support from a math coach who provided 5-min of individual assistance to each student.

The school reported approximately 34% of the students were having two to five office discipline referrals (ODRs) during the first two semesters of the current school year, and 5% of the students were having six or more ODRs. The school had been implementing school-wide universal supports within PBIS for 6 years. Data from the recent academic year showed their Benchmarks of Quality (BoQ; Cohen et al., 2007) score was 93%, indicating a high degree of fidelity in implementing school-wide universal supports.

2. Measurement

1) Academic Engagement and Disruptive Behavior

The dependent variables were class-wide academic engagement and disruptive behavior. Academic engagement was defined as students demonstrating active participation and focus, which included engaging in any of the following behaviors: (a) maintaining visual orientation toward instructional

materials or teacher while remaining in the assigned area with head oriented towards the materials or teacher, (b) raising hands, (c) answering questions, and (d) staying engaged in assigned work or tasks (e.g., looking through books, completing worksheets independently or with peers, engaging with instructional materials).

Disruptive behavior was defined according to classroom rules, which included engaging in any of the following behaviors: (a) off-task behavior (e.g., head down on desk, playing with objects unrelated to given task, or invading others' personal space), (b) disrespectful actions (e.g., yelling at others, poking others in the arm, running hand on back, taking the property of others), (c) interfering with learning activities (e.g., engaging in vocalizations unrelated to activity, entering other parts of the room without permission, tapping pencil on desk, crawling on floor), (d) disregarding instructions (e.g., refusing task), and (e) unsafe behavior (e.g., throwing objects, hitting others with objects or hands, deliberately falling to ground from standing position, rocking back and forth on chair with both feet off the ground, jumping on desk or table, kicking objects).

(1) Direct Observation

The researcher (first author) and two research assistants collected direct observational data 3 to 4 times per week. They recorded target behaviors during the initial 20-30 min (average, 29 min) of targeted time periods, using an electronic timer on a smartphone. The observation duration varied within and across classrooms. Academic engagement was measured using the planned activity check (PLACHECK) procedure, a variation of momentary time sampling designed to measure group behavior (Cooper et al., 2021). At the end of each 3-min interval, the observers systematically scanned the classroom in a predetermined order and counted the number of students engaged in academic activities. The percentage of academically engaged students was calculated by dividing the number of students engaged by the total number of students present at each planned check. The average level of appropriate behavior for each session was determined by summing the percentage of students engaged across checks and dividing by the number of checks.

Disruptive behavior was measured using a 15-s partial-interval recording system. Observers noted the occurrence of the disruptive behavior if it was exhibited by any student in the class during any portion of an interval. To determine the overall level of class-wide disruptive behavior, the percentage of intervals where disruptive behavior occurred was calculated. Academic engagement and disruptive behavior were not mutually exclusive, meaning that both behaviors could be observed simultaneously

during an interval.

(2) Teacher-Completed Behavior Rating Scale

To supplement direct observational data, teachers collected behavior rating scale (BRS) data on the class-wide disruption and academic engagement. The BRS was created using the Individualized Behavior Rating Scale Tool development guidelines (Iovannone et al., 2014) and designed to estimate the number of students engaging in the target behaviors. The BRS for disruption used a 6-point Likert-type scale with scores corresponding to the number of students engaging in disruptive behavior. Score 0-2 students was set at anchor point 1 (best day; least problem day); 2-4 students at 2; 4-6 students at 3; 6-8 students at 4; 8-10 students at 5; and 10+ students at 6 (worst day).

The BRS for academic engagement was scored inversely, with 1 representing the worst day and 4 representing the best day. At the conclusion of each instructional time, teachers marked the number of students they believed to be disruptive for the majority of the instructional time according to the operational definitions. They then circled the corresponding rating score for both disruptive behavior and academic engagement. After completing multiple sessions, teachers connected the circled scores to create a line graph, which facilitated visual interpretation of class-wide behavioral performance trends and data-based decision-making.

2) Implementation Fidelity

To objectively assess the teachers' adherence to each treatment component, an independent observer, a research assistant observed teacher implementation of the intervention while completing a 20-item implementation fidelity checklist with a yes/no format during 100% of the intervention sessions. The number of components implemented correctly was divided by the total number of components (10 components) and then multiplied by 100 to determine the percentage of implementation fidelity. Class 1 teacher's average fidelity was 92% (range, 70%-100%). Both Class 2 and Class 3 teachers demonstrated 100% fidelity throughout the intervention phase.

3) Interobserver Agreement (IOA)

The researcher and a research assistant simultaneously and independently collected direct observational data on class-wide target behaviors for 33% to 40% baseline and intervention sessions across classes to assess IOA on the collected data. The researcher trained two undergraduate students

in Applied Behavior Analysis to serve as research assistants, using YouTube videos of classroom students engaging in disruptive behavior similar to the students in the participating classes. The research assistants were required to achieve 90% or higher agreement on practice data training before collecting data.

For the PLACHECK observations, IOA was calculated by dividing the smaller number of students observed by the larger number of students observed for each check, summing these ratios, and dividing by the total number of checks. For the interval recording of disruptive behavior, IOA was calculated by dividing the number of intervals with agreement by the total intervals, then multiplying by 100%. Across phases, the average IOA was 94% in Class 1, 92% in Class 2, and 97% in Class 3 for academic engagement. For disruptive behavior, the average IOA was 89% in Class 1, 83% in Class 2, and 91% in Class 3. Across all behaviors, phases, and classes, IOA ranged from 83% to 100%, with the exception of one baseline or intervention session in each class where IOA dropped to 62%, 68%, or 78%, respectively. For teacher implementation fidelity, IOA was assessed on 100% of implementation sessions, calculated by dividing the number of agreed steps by the total number of steps, then multiplying by 100%. The IOA for implementation fidelity was 100% across all sessions for all classes.

4) Social validity

Teachers were asked to complete a modified, 15-item Intervention Rating Profile-15 rating scale (IRP-15; Martens et al., 1985) at the end of the intervention to assess the social validity of their chosen GC intervention. The items were assessed using a 6-point Likert-type scale to indicate whether the intervention was acceptable, effective, and efficient. The IRP-15 was slightly modified by changing the definitions from individual children to the whole class and describing the GC intervention.

3. Experimental Design and Procedures

The outcome of the GC intervention was evaluated using a concurrent multiple-baseline design across classrooms. Before collecting baseline data, the researcher and each classroom teacher collaboratively determined the definitions of disruptive behavior and academic engagement. The researcher then provided each teacher a10-min training on how to complete the BRS and evaluate

changes in level, trend, and variability of class behavior. The teachers received an instruction sheet with information detailing how to complete the BRS, to which they could refer at any time.

1) Baseline

During this phase, teachers conducted class as usual, with students receiving universal support. At the beginning of the class, the teachers reviewed classroom expectations and rules posted on classroom walls, which were aligned with to the school-wide expectations and rules. All teachers provided behavior-specific praise and school-wide tokens to students based on the students' appropriate behaviors. All teachers referred to school-wide expectations and rules when there were instances of problem behaviors. For severe problem behavior, the teachers referred the students to the office. The teachers also used a level system where students moved clips up and down on a chart with either various colors based on each student's appropriate or problem behavior during class. Baseline data were used for each classroom to assist in determining goals for the level of disruptive behavior and academic engagement during the intervention phase.

2) Selection of Reinforcers

The classroom teachers conducted a brief preference assessment using a 4-question Likert-type survey, identifying reinforcers they felt were appropriate for the class as a whole compared to individual students, and created a list of mystery motivators (MMs). The teachers showed this list to the students, giving them the opportunity to select three items or activities that they most preferred. The items or activities selected by at least 25% of students were chosen as MMs. The preferred items were relatively inexpensive or free, except a few edibles that were provided by the study team, and the preferred activities were readily available in the school. The MMs included candies, cookies, chips, notebooks, mechanical pencils, playing a game, playing in the playground, free play, time with animals, and time with stuffed animals. The amount of edibles or tangibles given to students was determined daily by the teacher, if applicable.

3) Teacher Training and Selection of a Preferred GC

The researcher provided a 45- to 80-min individual teacher training session on implementing GCs, as the teachers had no prior experience with them. Training consisted of general instruction on how to implement the different GCs (randomized, independent, dependent, and interdependent), along

with some background information on the contingencies, including benefits and issues. Teacher training included a PowerPoint presentation with written materials and brief YouTube videos of GCs. At the end of the training, the teachers chose not to implement the dependent contingency due to the possibility of stigmatizing individual students. The teachers completed a 4-item preference assessment survey for each of the three remaining GCs. The survey included a description of the GC and questions regarding whether the contingency type might be more acceptable, effective, and a good fit for the classroom students. The researcher informed the teachers that once they selected their preferred type, they could not alter the GC. Based on their responses to the preference survey, the teachers chose their GC type: independent (Class 1 and Class 2) and randomized (Class 3).

The teachers examined the baseline BRS data and set goals for disruptive behavior reduction levels and academic engagement increase levels based on mean baseline levels. The reinforcement criteria for preferred GC varied. The selected criteria were alternating between 4 and 8 rule violations for Class 1 (independent), between 2 and 8 rule violations for Class 2 (independent), and between 5 and 7 rule violations for independent GC and 7-9 for interdependent GC for Class 3 (randomized). The criterion for the independent contingency looked like “No more than 5 Xs”. The criterion for interdependent contingency looked like, “Class total Xs” or “Check marks less than 7”.

4) GC Implementation with Data-Based Decision-Making

The researcher and each teacher provided a 10-min GC training session to their students before the intervention implementation. The training involved reviewing class expectations, rules, and examples of rule-following and rule-breaking behaviors. The teacher-chosen GC was reviewed, and the list of MMs was introduced. The training stressed the importance of encouraging peers to be kind and following class rules to earn the reward for the day. Teachers implemented their chosen, preferred GC with data-based decision-making procedures. The teachers had two or three selection boxes for randomized elements. These boxes had the following labels (depending on the GC chosen): reinforcement criteria, MM, and group reward type. Strips of paper that identified all choices for each element were placed in the box whereby each choice was drawn randomly (e.g., two types of GCs, four criteria for reinforcement, and five MMs). The teachers used a one-page simplified chart guided how to implement the procedures and included a brief script to read to students, identification of boxes from which to draw for each contingency type, and how to provide and limit access to the MMs. The teachers began the instructional activity by reading the script and explaining

how students would have the opportunity to earn an MM. The teachers then explained the expectations and the range of criteria for disruptive behavior to the students but did not identify the MM. After the script was read to the students, the teachers continued teaching her planned lessons.

When the independent contingency was implemented, students received a check mark on the chart beside their name each time they engaged in disruptive behavior. The teachers could say something like "Johnny has earned a check because he talked out." At the end of the instructional time, the teachers announced the end of implementation, selected the criterion and MM from the respective boxes, and compared the criterion to the number of check marks next to each student's name. If students met the criterion or had less than the criterion number of checks, they received praise for following school expectations and classroom rules along with the MM. If students exceeded the criterion, the teachers stated, "Well you weren't able to earn the reward this time, but you can have another chance to earn it later."

When implementing the interdependent contingency, students who engaged in disruptive behavior received check marks in the same manner as the independent contingency. The teacher recorded check marks for all students in a manner that was visible to the students. After the instructional activity, the teachers announced the end of implementation, selected a criterion from the box, and compared it with the number of check marks for the class as a whole. If the class met the criterion for reinforcement, the teachers then selected an MM from the box, praised the students for being in alignment with school expectations and classroom rules, and gave the students the MM. If the class did not achieve the criterion, the teachers encouraged them to try again next time.

During implementation of the randomized contingency, students could earn the MM based on the behavior of each child individually (independent contingency) or the class as a whole (interdependent contingency). Prior to the instructional activity, the teacher announced to the class the expectations and the range of criteria for disruptive behavior but did not reveal the GC type; the students did not know how access to reinforcement would be determined. At the end of the instructional period, the teacher selected the GC type from the GC Type box and completed the procedures for determining access to the reinforcement based on which contingency type was being implemented.

The teachers also collected BRS data at the end of each session. They continued to implement their preferred contingency without any modification to their procedures when their BRS demonstrated continued improvement towards class goals. When the BRS data did not demonstrate continued improvement or ran counter to preferred change after a few sessions, the teachers

implemented the data-based decision-making procedures. The teachers reviewed the previous week's BRS data and identified problems that might have affected the undesirable changes in the target behaviors. When the BRS data showed increases in disruption and decreases in academic engagement, they changed criteria for gaining reinforcers (Class 1 and Class 3), increased the length of the implementation session (Class 2), or increased student buy-in by having them perform the picking from the MM box and Criteria for Reward box (Classes 2 and 3). The teachers informed the researcher that they would start changing components and, on occasion, collaborated with the author during the data-based decision-making to identify the components that required modification. However, the teachers made modifications to their GC procedures on their own based on their data rather than requesting additional training or participating in the problem-solving process with the author.

5) Follow-up

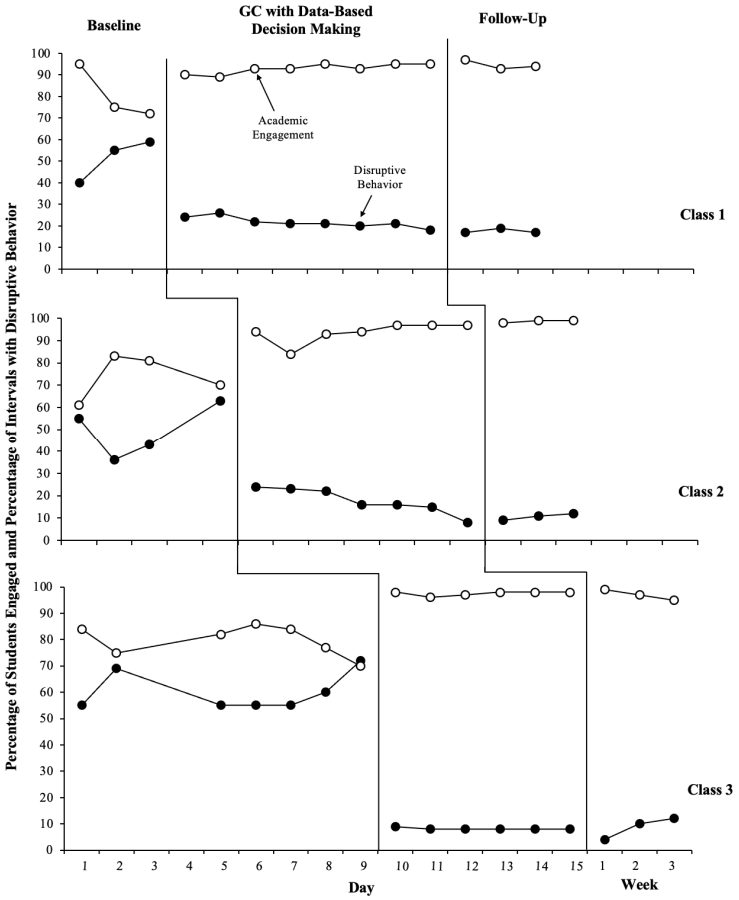
Beginning two weeks after the intervention, three probes were conducted once a week. Teachers were no longer being asked to implement the GC intervention following termination of the intervention phase. However, all three classroom teachers reported that they chose to continue implementing their preferred GC intervention during follow-up.

Ⅲ. Results

1. Direct Observational Data

1) Academic Engagement

<Figure 1> depicts class-wide academic engagement data. Data indicated that the intervention increased academic engagement in all three classes. Academic engagement increased immediately upon implementation of the intervention and remained stable over the course of intervention. In baseline, the mean academic engagement (percentage of students engaged) for Class 1 was 80.7% (range, 72%-95%), demonstrating a decreasing trend. During intervention, academic engagement increased to a mean of 92.9% (range, 89%-95%). The academic engagement level remained high across sessions during intervention. Class 2 showed a similar pattern. Academic engagement increased from 76.2%



<Figure 1> Class-wide academic engagement and disruptive behavior across classes and phases

(range, 61%-83%) in baseline to 95% (range 84%-99%) in intervention. Compared to baseline, data were stable during intervention, remaining at over 93% after two sessions. Class 3 also demonstrated an immediate change in academic engagement when the intervention was implemented. In baseline, academic engagement averaged 79.7% (range: 70%-86%), with a decreasing trend toward the end. During intervention, it increased to a mean of 97.4% (range 95%-99%) with a stable pattern.

2) Disruptive Behavior

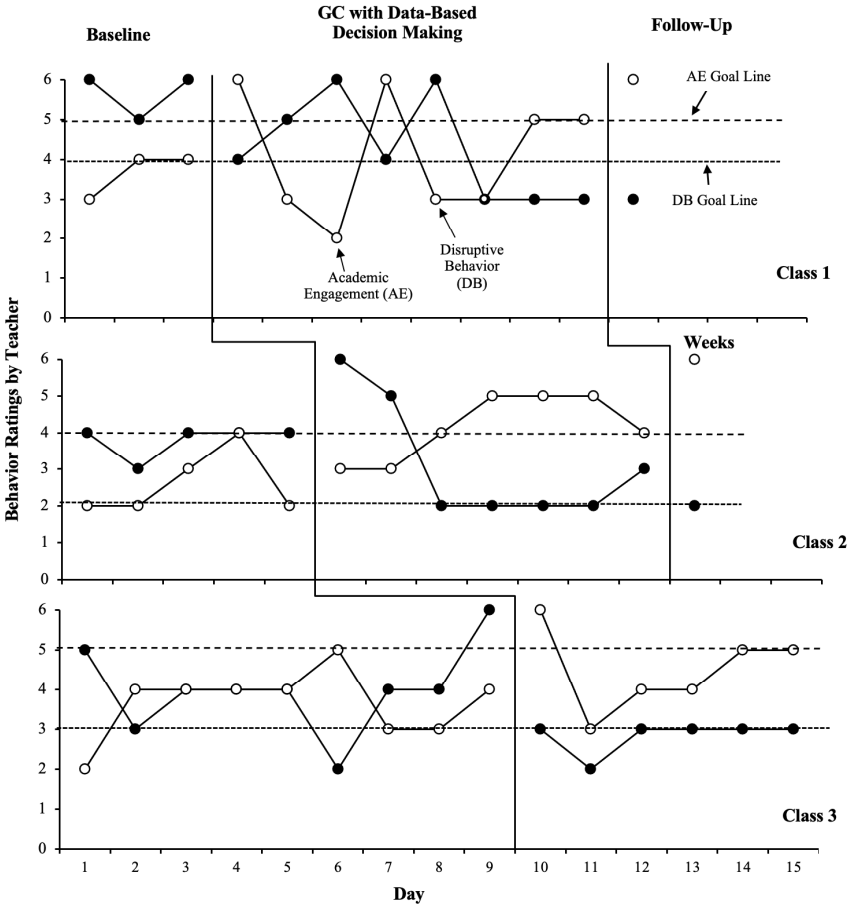
<Figure 1> also depicts class-wide disruptive behavior across phases. The data indicated the teacher-preferred GCs implemented with data-based decision-making resulted in decreased disruptive behavior in all three classes. For Class 1, disruptive behavior occurred during a mean of 51.3% of

intervals (range, 40%-59%) with an increasing trend in baseline. Implementing the teacher-preferred GC (independent) resulted in a decrease in disruptive behavior to a mean of 21.6% (range, 18%-26%), a period during which the level of disruptive behavior remained low and trended slightly downward. For Class 2, disruptive behavior occurred during a mean of 48% (range, 36%-63%) with an increasing trend in baseline. The introduction of the teacher-preferred GC (independent) resulted in an immediate decrease in disruptive behavior to a mean of 17.9% (range, 8% – 24%), and the data demonstrated a stable pattern. For Class 3, disruptive behavior occurred a mean of 60.3% of intervals in baseline (range, 55%-72%). Baseline data were somewhat variable during initial sessions but became stable and showed an increasing trend toward the end of baseline. When the GC (randomized) was implemented, disruptive behavior decreased to a mean of 8.2% (range, 8%-9%).

2. Behavior Rating Scale Data

<Figure 2> presents the teacher-collected BRS data. The BRS data indicated that the teachers' observed levels of disruptive behavior decreased, and academic engagement increased in all classes as a result of implementing the intervention. The teachers' ratings for disruptive behavior averaged 5.7, 3.8, and 4.0 in baseline and 4.1, 3.0, and 2.8 in intervention for Class 1, Class 2, and Class 3, respectively. For academic engagement, the ratings averaged 3.7, 2.6, and 3.7 in baseline and 4.3, 4.4, and 4.5 in intervention for Class 1, Class 2, and Class 3, respectively.

While the direct observational data consistently demonstrated lower rates for problem behavior and higher rates for academic engagement in intervention compared to the baseline, the Class 1 teacher's ratings for both behaviors were variable during intervention. The class met the goals in 63% of the sessions for problem behavior and 50% for academic engagement. For Class 2, ratings for disruptive behavior in intervention were initially higher than those in baseline, although the direct observational data indicated an immediate decrease as the intervention was implemented. However, the ratings for disruptive behavior in later intervention sessions were consistently lower than those at baseline. The Class 2 ratings for academic engagement in intervention initially showed an increasing trend, then stabilized during the remaining sessions with the exception of the last session. The class met the goals in 73% of the sessions for both target behaviors. The Class 3 ratings for disruptive behavior were consistently low (2 or 3 out of 6) in intervention, whereas the rating for academic engagement was high (6) in session 1 but decreased to 3 in session 2. However, the data on academic



<Figure 2> Behavior rating scores by teachers for academic engagement and disruptive behavior

engagement demonstrated an increasing trend in later intervention sessions. The class met the goals in 83% of the session for problem behavior and 50% for academic engagement.

3. Maintenance

Follow-up observations conducted 2 weeks after the intervention showed the maintenance of improved class-wide behavior for all classes. Class 3 showed a slightly increasing trend for problem behavior and a decreasing trend for academic engagement; however, the levels were much better than the baseline levels. Class 1 and Class 2 teachers completed one follow-up probe. Their ratings also indicated maintenance of changes in both behaviors, meeting or exceeding the set goals.

4. Social Validity

The IRP-15 completed by three teachers revealed that the teacher-preferred GC with data-based decision making was rated as highly acceptable and effective. Teacher ratings averaged 5.1 out of 6, ranging from 4 to 6 across items, indicating a high level of acceptability and satisfaction with the intervention.

Ⅳ. Discussion

This study examined the impact of a teacher-preferred GC intervention, integrated with data-based decision-making, on class-wide disruptive behavior and academic engagement across three classrooms in an urban public elementary school. Teachers implemented the intervention with high levels of fidelity throughout the intervention. Teachers implemented the intervention with consistently high fidelity throughout the intervention. The intervention immediately produced reductions in disruptive behavior and corresponding increases in academic engagement across all three classrooms. These behavioral improvements were sustained at 2-week follow-up observations. The classroom teachers conducted data-based decision-making using BRS data and adjusted the contingency criteria, session length, or student involvement in selecting criteria or MMs without guidance from the researcher. No additional teacher training was necessary during intervention. Teachers unanimously reported that the intervention was highly acceptable and effective, and they liked implementing GCs.

1. Implications

The results of this study support previous research in that GCs are effective in increasing academic engagement or on-task behavior and decreasing problem behavior during classroom sessions (Alric et al., 2007; Bohan et al., 2022; Hirsch et al., 2016; Ling et al., 2011) and that a teacher-preferred GC can enhance class-wide student behavioral outcomes in elementary schools (Ennis et al., 2016). In Ennis et al.'s study, teachers used four types of GCs: independent, dependent, interdependent, and randomized. In the first phase of the intervention, they switched between the four types, and in the second phase, they used their favorite type of GC. Ennis et al. suggested that although all the GCs

were equally effective in improving class-wide behaviors, teachers' preference on the types of GCs could vary depending on their instructional practices and classroom's ecology. A preference assessment on GC types may promote buy-in from teachers, which is critical to implementing the intervention with fidelity and to improving student behavior and classroom ecology (Ennis et al., 2016).

In comparing independent and interdependent GCs within the Good Behavior Game (GBG, a variation of the interdependent GC) with four students with severe emotional and behavioral disorders, aged 9 – 10 years old, Groves and Austin (2017) reported that both versions of GBG were equally effective in decreasing problem behaviors for all students. Following the comparison of the two versions of the GBG, the researchers implemented each student's preferred version in the subsequent phase. The researchers suggested that the interdependent GC was not an essential component of the GBG; instead, teacher or student preferences may be a key factor influencing the successful implementation of the GBG to affect student behavior.

Donaldson et al. (2018) examined the use of GBG to reduce disruptive behavior in three kindergarten classes and one first-grade class. The researchers examined teacher preference for who would implement the GBG (teacher vs. student) following a comparison of experimenter-led, student-led, and teacher-led GBG. The researchers found that the GBG was effective in reducing student disruptive behavior regardless of who implemented it and also suggested that teacher preference should be taken into account by involving teachers in determining which and how class-wide interventions should be implemented, considering that teacher preferences for who would implement the GBG varied across teachers.

The high social validity of the intervention in the current study indicates that the three teachers valued the GCs they preferred and could implement with fidelity. Furthermore, with minimal training on the BRS and data-based decision-making, the teachers efficiently and effectively modified their GC implementation procedures based on collected BRS data. The results demonstrated large intervention effects; no intervention data points for disruptive behavior overlapped with the baseline data points in any of the three classrooms, and no intervention data points for academic engagement overlapped with the baseline data points in two classrooms. Although direct observational data demonstrated stability during the initial intervention sessions, with higher rates of academic engagement and lower rates of disruptive behavior compared to baseline, teachers' ratings on both behaviors were less desirable than their expectations or goals, which led to making modifications to their procedures. One reason that there were differences in level and variability of data between the

two data sources might be that the direct observational data were collected during the first 20-30 min of their activity period in the case of classes that lasted 30 min or longer, whereas the teachers' ratings were based on their entire activity period and that disruptive behavior might have occurred at higher rates during some days.

One factor that affected the successful outcomes in this study might be the development of operational definitions for disruptive behavior and criteria for contingencies by linking school-wide expectations and classroom rules. It was observed that the teachers reviewed the expectations and rules every time that they were implementing the intervention. In Ennis et al. (2016), teachers tended to bypass reviewing the classroom rules and expectations when they were implementing their preferred contingency every day, which might have resulted in relatively small changes in appropriate behavior across classrooms in their study.

This study extends existing literature by demonstrating that teachers could effectively implement the GC procedures with fidelity after brief training, which included instructions, modeling, and a simplified individualized instruction sheet. As indicated in previous studies, GCs demonstrate practicality and adaptability, making them feasible for classroom settings due to their contextual fit, allowing teachers to implement them easily (Ennis et al., 2016; Skinner et al., 2009; McIntosh et al., 2010). In the current study, the teachers' successful implementation of GCs with data-based decision-making, without extensive external expert training, may be attributed to their prior exposure to data-driven approaches through multi-tiered systems of supports (MTSS) or school-wide positive behavioral interventions and supports (PBIS). These implementation frameworks prioritize tiered interventions that address diverse student needs, emphasize problem-solving, and utilize data-based decision-making to improve students' academic and behavioral outcomes (McIntosh & Goodman, 2016).

2. Limitations and Future Directions

The findings from this study can be limited by its small sample size; only three classes from three grades participated in the study. The teachers' selection of contingencies might not accurately represent the population of elementary school general education teachers. This small size was due to the time commitment for implementation of GC conditions and the length of time it took to obtain parental permission forms. Another limitation could have arisen from the data collection methods;

different measurement systems were used for disruptive behavior and academic engagement. A partial interval recording system was used to measure disruptive behavior, as these behaviors were brief and discrete, whereas PLACHECK was selected for academic engagement to accommodate its continuous nature without a clear beginning or end.

Overall, IOA was acceptable across classes, behaviors, and phases; however, a few sessions had low IOA, particularly for disruptive behavior. This may have been influenced by using different recording systems for data collection and the challenge of simultaneously observing the behavior of 18 individual students. In addition, certain disruptive behaviors, such as talking to others and placing objects on another's desk, were more challenging to observe due to their subtlety and dependence on the observer's position relative to the student. Measuring academic engagement was also difficult during transitions within the classroom, particularly when teachers did not clarify acceptable or unacceptable behaviors during these times. Furthermore, inconsistencies in teacher feedback regarding behaviors such as keeping heads down while working may have led to variations in recording disruptive behaviors and academic engagement.

An additional limitation is with follow-up data. The study collected only three weekly follow-up data points during which the teacher implemented the intervention. Consequently, the limited data collection makes it challenging to determine whether the GC intervention with data-based decision-making can effectively promote long-term maintenance of behavioral improvements after the intervention has been terminated. Further research evaluating long-term maintenance assessment would increase confidence in the findings. Another limitation pertains to the variations in marking students' names on a chart when recording rule violations during the implementation of contingencies. Teachers varied in their frequency and immediacy of providing private, public, or no feedback, and in their recording of disruptive behavior. These differences, along with different grade levels, complicate direct comparisons between classes. However, such diversity likely enhanced the contextual fit and intervention acceptability across different classroom environments..

3. Conclusion

In this study, the group reinforcement procedures provided teachers with a structured approach to address disruptive behavior while implementing school-wide universal supports in the classroom. These additional supports facilitated student success by reinforcing positive behaviors and adherence to

classroom expectations and rules. The findings indicate that integrating GC procedures into elementary classrooms can significantly improve classroom management. The results validate the hypothesis that GCs are effective and highlight the value of data-based decision-making in the classroom. Before implementing a GC, teachers should consider conducting preference assessments to identify the most appropriate types of GCs, as teacher-preferred GCs can optimize student behavioral outcomes and facilitate more sustainable classroom management strategies. Furthermore, additional professional development focused on creating and utilizing behavior rating scales could enhance teachers' ability to implement data-based interventions more effectively.

References

Alric J. M., Bray M. A., Kehle T. J., Chafouleas S. M., & Theodore L. A. (2007). A comparison of independent, interdependent, and dependent group contingencies with randomized reinforcers to increase reading fluency. *Canadian Journal of School Psychology, 22*(1), 81-93. https://doi.org/10.1177/0829573507301254

Benazzi, L., Horner, R., & Good, R. (2006). Effects of behavior support team composition on the technical adequacy and contextual fit of behavior support plans. *The Journal of Special Education, 40*(3), 160-170. https://doi.org/10.1177/00224669060400030401

Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 9-22. https://doi.org/10.1177/003172170408600105

Bohan C., McDowell C., Smyth S. (2022). Does the immediacy of feedback matter in game-based classroom management? Analysis of the caught being good game with adolescent students. *Journal of Positive Behavior Interventions*, *24*(3), 208-221. https://doi.org/10.1177/10983007211068534

Brennan, C., Deegan, A., & Smyth, S. (2024). A systematic review and narrative synthesis of single-case group contingency interventions targeting prosocial and antisocial behavior in school children. *Journal of Positive Behavior Interventions*. Advance online publication. https://doi.org/10.1177/10983007241235

Chafouleas, S.M. (2011). Direct Behavior Rating: A review of the issues and research in its development. *Education and Treatment of Children, 34*(4), 575-591. https://doi.org/10.1353/etc.2011.0034

Chafouleas, S. M., Riley-Tillman, T. C., & Christ, T. J. (2009). Direct Behavior Rating (DBR): An emerging method for assessing social behavior within a tiered intervention system. *Assessment for*

Effective Intervention, 34(4), 195-200. https://doi.org/10.1177/1534508409340391

Cihak, D. F., Kirk, E. R., & Boon, R. T. (2009). Effects of class-wide positive peer "tootling" to reduce the disruptive classroom behaviors of elementary students with and without disabilities. *Journal of Behavioral Education, 18,* 267-278. https://doi.org/.1007/s10864-009-9091-8

Cohen, R., Kincaid, D., & Childs, K. (2007). Measuring school-wide positive behavior support implementation: Development and validation of the Benchmarks of Quality. *Journal of Positive Behavior Interventions, 9*(4), 203-213. https://doi.org/10.1177/10983007070090040301

Cooper, J. O., Heron, T. E., & Heward, W. L. (2021) *Applied behavior analysis* (3rd.). Upper Saddle River.

Donaldson, J. M., Matter, A. L. & Wiskow, K. M. (2018). Feasibility of and teacher preference for student-led implementation of the good behavior game in early elementary classrooms. *Journal of Applied Behavior Analysis, 51*(1), 118-129. https://doi.org/10.1002/jaba.432

Ennis, C. R., Blair, K. C., & George, H. P. (2016). An evaluation of group contingency interventions and the role of teacher preference. *Journal of Positive Behavior Interventions, 18*(1), 17-28. https://doi.org/10.1177/1098300715577663

Ford, S., Blair, K. C, Iovannone, R., & Kwak, D. (2024). Implementation of the Prevent-Teach-Reinforce model for elementary school students needing intensive behavior intervention. *Behavioral Sciences, 14*(2), 93. https://doi.org/10.3390/bs14020093

Groves, E. A., & Austin, J. L. (2017). An evaluation of interdependent and independent group contingencies during the good behavior game. *Journal of Applied Behavior Analysis*, *50*(3), 552-566. https://doi.org/10.1002/jaba.393

Heering, P. W., & Wilder, D. A. (2006). The use of dependent group contingencies to increase on-task behavior in two general education classrooms. *Education and Treatment of Children, 29*(3), 459-467.

Hirsch, S. E., Healy, S., Judge, J. P., & Lloyd, J. W., (2016). Effects of an interdependent group contingency on engagement in psychical education. *Journal of Applied Behavior Analysis, 49*(4), 957-979. https://doi.org/10.1002/jaba.328

Hoover, J., & Patton, J. R. (2004). Differentiating standards-based education for students with diverse needs. *Remedial and Special Education, 25*(2), 74-78. https://doi.org/10.1177/07419325040250020101

Hulac, D., & Benson, N. (2010). The use of group contingencies for preventing and managing disruptive behaviors. *Intervention in School and Clinic, 45*(4), 257-262. https://doi.org/10.1177/1053451209353

Ingram, D., Louis, K., & Schroeder, R. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record, 106*(6), 1258-1287. https://doi.org/10.1111/j.1467-9620.2004.00379.x

Iovannone, R., Greenabum, P., Wang, W., Dunlap, G., & Kincaid, D. (2014). Interrater agreement of the individualized behavior rating scale tool. *Assessment for Effective Interventions, 39*(4), 195-207. https://doi.org/10.1177/1534508413488414

Kelshaw-Levering, K., Sterling-Turner, H., Henry, J., & Skinner, C. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the Schools, 37*(3), 523-533. https://doi.org/10.1002/1520-6807(200011)37:6<523::AID-PITS5>3.0.CO;2-W

Kilgus, S. (2013). Use of direct behavior ratings as the foundation of tier 2 service delivery. *Journal of Curriculum and Instruction, 7*(1), 79-99. https://doi.org/10.3776/joci.2013.v7n1p79-99

Ling, S., Hawkins, R., & Weber, D. (2011). Effects of a class-wide interdependent group contingency designed to improve the behavior of an at-risk student. *Journal of Behavioral Education, 20*(2), 103-116. https://doi.org/10.1007/s10864-011-9125-x

Little, S., Akin-Little, A., & Newman-Eig, L. (2010). Effects on homework completion and accuracy of varied and constant reinforcement within an interdependent group contingency system. *Journal of Applied School Psychology, 26*(2), 115-131. https://doi.org/10.1080/15377900903471989

Maggin, D. M., Pustejovsky, J. E., & Johnson, A. H. (2017). A meta-analysis of school-based group contingency interventions for students with challenging behavior: An update. *Remedial and Special Education, 38*(6), 353-370. https://doi.org/10.1177/0741932517716900

Martens, B. J., Witt, J. C., Elliot, S. N, & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice, 16*(2), 191-198. https://doi.org/10.1037/0735-7028.16.2.191

McIntosh, K., Filter, K., Bennett, J., Ryan, C., & Sugai, G. (2010). Principles of sustainable prevention: Designing scale-up of school-wide positive behavior support to promote durable systems. *Psychology in the Schools, 47*(1), 5-21.

McIntosh, K., & Goodman, S. (2016). *Integrated multi-tiered systems of support: Blending TRI and PBIS*. The Gilford Press.

McKissick, C., Hawkins, R., Lentz, F., Hailley, J., & McGuire, S. (2010). Randomizing multiple contingency components to decrease disruptive behaviors and increase student engagement in an urban second-grade classroom. *Psychology in the Schools, 47*(9), 944-959. https://doi.org/10.1002/pits.20516

McNaughton, S., Lai, M., & Hsiao, S. (2012). Testing the effectiveness of an intervention model based on data use: A replication series across clusters of schools. *School Effectiveness and School Improvement, 23*(2), 203-228. https://doi.org/10.1080/09243453.2011.652126

Skinner, C., Skinner, A., & Burton, B. (2009). Applying group-oriented contingencies in classrooms. In

K. A. Akin-Little, S. G., Little, M. Bray, & T. Kehle (Eds.), *Behavioral interventions in schools: Evidence-based positive strategies* (pp. 157-170): American Psychological Association. https://doi.org/10.1037/11886-010

Skinner, C. H., Skinner, A. L., & Sterling-Turner, H. E. (2002). Best practices in utilizing group contingencies for intervention and prevention. In A. Thomas & S. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 817-830). National Association of School Psychologists.

Skinner, C. H., Williams, R. L., & Neddenriep, C. E. (2004). Using interdependent group-oriented reinforcement to enhance academic performance in general education classrooms. *School Psychology Review, 33*(3), 384-397.

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development (2009). *Implementing data-informed decision making in schools: Teacher access, supports, and use.*

Wayman, J. C. (2005). Involving teachers in data-driven decision making: Using computer data systems to support teacher inquiry and reflection. *Journal of Education for Students Placed at Risk. 10*(3), 295-308. https://doi.org/10.1207/s15327671espr1003_5

국문초록

교사가 선호한 집단강화와 자료기반 의사결정의 사용: 학급차원 행동 개선

Fernando Herrera, M.A., BCBA (Medical University of the Americas)
Kwang-Sun Cho Blair (조광순), Ph.D., BCBA-D (University of South Florida)

이 연구는 학급 차원의 행동 개선을 하고자 초등학교 세 학급에서 교사가 선호한 집단강화를 자료기반 의사결정과 함께 실행하여 그 효과를 알아본 것이다. 중재과정에서 세 학급의 교사들은 독립적 집단강화 또는 무작위로 선택한 집단강화를 자신들이 선호하는 집단강화로 선택하여 실행하였으며, 행동평가 척도를 사용해서 수집한 행동자료를 토대로 집단강화 기준, 세션 길이, 강화 유형에 관한 결정을 하였다. 이러한 자료기반 의사결정과 함께 실행한 집단강화가 학급차원 교실 행동에 미치는 영향을 알아보기 위해 중다기초선 설계를 사용하였다. 연구 결과, 교사가 선호한 집단강화를 자료기반 의사결정 절차와 함께 실행할 경우 수업방해 행동을 크게 감소시키고 수업참여행동은 적절히 증가시킬 수 있음을 모든 세 학급에서 보여주었다. 또한 이러한 학급차원 행동 변화는 2주 후 추적 조사에서도 유지되고 있는 것을 보여 주었다.

주제어 : 집단강화, 학급차원, 교사 선호성, 자료기반 의사결정, 수업방해행동

게재 신청일 : 2024. 11. 21
수정 제출일 : 2024. 12. 10
게재 확정일 : 2024. 12. 10

* Corresponding Author : kwangsun@usf.edu