

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
1–14

© Hammill Institute on Disabilities 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10983007211068534

jpbj.sagepub.com



Clare Bohan, PhD¹, Claire McDowell, PhD²,
and Sinéad Smyth, PhD¹

Abstract

This study evaluated the Caught Being Good Game (CBGG) across two adolescent student populations, maintaining a focus on the provision of feedback during the game. The CBGG, a variation of the group contingency intervention, the Good Behavior Game (GBG), is a classroom management intervention that involves the provision of points to teams of students who follow class rules. Feedback was manipulated during the game to ascertain whether immediate visual feedback was always necessary. The CBGG was presented with and without immediate visual feedback across phases, using a multiple treatment reversal design. Intervention conditions were counterbalanced across two classrooms of mainstream adolescent students. Data were collected on academically engaged and disruptive behavior. The CBGG was generally effective in targeting these behaviors in both classrooms, with some differential effects apparent for CBGG versions across classrooms. This provides further support for the use of the CBGG as a positive classroom management technique and as an alternative to the classic GBG. The findings also suggest that teachers may choose whether to use feedback or not during the CBGG, which may save them time and increase buy-in by incorporating an opportunity for some autonomy in game implementation.

Keywords

adolescent, academic engagement, Caught Being Good Game, Tier I positive behavior intervention

Introduction

The Good Behavior Game (GBG) is a group contingency intervention for classroom settings, which has been successful in targeting disruptive behaviors and academic engagement, as demonstrated by more than 50 years of empirical research (Barrish et al., 1969; Bowman-Perrott et al., 2016). In its most basic form, the game involves the development of a clear set of class rules, division of a class group into teams, and provision of marks or fouls to teams when a team member breaks a rule (i.e., an interdependent group contingency; Litow & Pumroy, 1975). Teams are eligible for prizes if they remain under a criterion of marks or fouls. Although this format has been effective in diverse classroom settings (e.g., Barrish et al., 1969; Dadakhodjaeva et al., 2020; Donaldson et al., 2017; Kleinman & Saigh, 2011; Mitchell et al., 2015; Nolan et al., 2014), researchers have attempted to delineate the essential features (Foley et al., 2019; Wiskow et al., 2019) and the game has been modified to focus on positive behavior in recent years (e.g., Wahl et al., 2016; Wright & McCurdy, 2012).

One of the most common modifications of the GBG involves the teacher providing points to teams of students for rule-following, as opposed to marks or fouls for rule-breaking. This modification has been called the “Caught Being Good Game” (CBGG) in recent publications (e.g., Bohan et al., 2021; Wahl et al., 2016; Wright & McCurdy, 2012) and may help to bring the GBG more in line with modern, positive behavior practices in schools. It targets similar behaviors to the GBG and has been effective in targeting academic engagement and disruptive behaviors across classrooms (e.g., Bohan et al., 2021; Wahl et al., 2016; Wright & McCurdy, 2012). It has also been effective

¹Dublin City University, Ireland

²Ulster University, Coleraine, UK

Corresponding Author:

Clare Bohan, School of Psychology, Dublin City University,
DCU Glasnevin Campus, H268, Dublin 9, Ireland.

Email: Clare.bohan22@mail.dcu.ie

Action Editor: Todd Haydon

in targeting novel behavior, such as mobile phone usage, when paired with an antecedent intervention (Hernan et al., 2019). This modification is important and has warranted a new game title because the behavioral principles underlying the GBG and CBGG fundamentally differ. The GBG incorporates positive punishment, in that students receive a mark for engaging in behavior that is inconsistent with class rules. This has been recognized by leading researchers in the field, who have stated that marks or fouls can serve as punishers resulting in decreases in target behaviors during the game (e.g., Joslyn et al., 2019; McKenna & Flower, 2014; Wright & McCurdy, 2012). The GBG can also be conceptualized as incorporating differential reinforcement of low rates of behavior (DRL), where teams are eligible for a prize if they maintain their marks below a certain criterion. Conversely, the central principle underlying the CBGG is positive reinforcement, where teams receive points for engaging in behavior that aligns with classroom rules. Researchers have recently conducted research comparing both versions. In these comparison studies, the GBG and CBGG have been found to be similarly effective in primary school classrooms (e.g., Wahl et al., 2016; Wright & McCurdy, 2012). The CBGG has also been investigated independently of the GBG and has been successfully applied in secondary school classrooms (e.g., Bohan et al., 2021; Dadakhodjaeva, 2017; Ford, 2017).

Evaluation of the CBGG and other modifications of the GBG, including studies that have aimed to define the essential features of the intervention, have dominated the group contingency literature in recent years. Foley et al. (2019) conducted a full component analysis of the GBG in a preschool classroom. The researchers evaluated components of the GBG and found that all major components were necessary before a decrease in disruptive behavior was apparent. After the partaking class had been exposed to the whole GBG package, the GBG without contingent reinforcement was effective in maintaining low levels of disruption. Wiskow et al. (2019) looked specifically at feedback during the GBG and also focused on preschool classrooms. The lowest rates of disruptive behavior were observed when vocal feedback or visual + vocal feedback was given in response to rule violations during the GBG. Only recently have authors begun to look at essential components of the CBGG as opposed to the GBG. For example, a recent study by Groves and Austin (2020) examined a known versus unknown criterion during the CBGG with a Welsh Year 4 class (aged 8–9 years). The authors demonstrated that the CBGG was effective in reducing disruptive behavior across several targeted students, whether the criterion was made known to them at the beginning of the game or kept a “mystery” until the end of the game. These potentially modifiable components of the GBG and CBGG are crucial to our understanding of the games’ applicability in the classroom and can have important time-saving implications.

Bohan et al. (2021) manipulated feedback as a variable component of the CBGG. The CBGG was implemented in two formats with an adolescent class group. These two formats were termed the CBGG-d and the CBGG-i. During the CBGG-d, points were recorded by the participating teacher privately and feedback on team progress toward a weekly goal was withheld until the end of the game period. During the CBGG-i, the teacher recorded points publicly on a scoreboard during the game, meaning students experienced consistent visual feedback throughout. Unique to the CBGG and GBG literature, prizes were administered weekly, rather than daily, during both the CBGG-d and the CBGG-i. The manipulation of reinforcement delivery was therefore two-fold; delayed feedback in the form of point delivery at the end of class only and prizes being available only weekly rather than daily. Bohan and colleagues reported that both versions of the CBGG applied were effective in targeting engagement and disruptive behavior but that the CBGG-d produced slightly more stable results. From a behavioral perspective, and in applied settings such as classrooms, delaying feedback and reinforcement is of great practical significance. Delayed feedback is potentially challenging to arrange, given the need to plan ahead; however, it allows for participants to exercise self-control, forgoing engaging in behavior yielding immediate reinforcement (e.g., engaging in disruptive behavior for peer attention) in favor of those which yield delayed reinforcement (e.g., engaging with class expectations to gain points that are revealed at the end of class and a weekly prize; Stromer et al., 2000). While Bohan et al.’s findings provide a useful first step in examining ways of streamlining the CBGG, the fact remains that the research design chosen allowed for the possibility of a sequence effect. Specifically, because the CBGG-d was applied first by Bohan et al. as part of an ABACABAC reversal design, there is the possibility that the CBGG-d was more effective, because it was the first version of the game that students were exposed to. Bohan et al. recognized this limitation and pointed toward the potential for future research to implement the game across two classrooms, counterbalancing conditions.

Study Purpose

The CBGG provides a positive classroom support mechanism, focusing solely on the principle of reinforcement. The popular GBG by contrast can be interpreted as using DRL and punishment. It is evident that, although the CBGG has garnered empirical interest in recent years, there are still gaps in the literature in terms of populations studied and component analyses. In general, applications of the GBG and CBGG with younger students (e.g., Groves & Austin, 2020; Tanol et al., 2010; Wahl et al., 2016; Wright & McCurdy, 2012) are more plentiful than those with older students. The first aim of this study is therefore to evaluate the CBGG with adolescent students

in an Irish mainstream classroom context, a population only previously evaluated by Bohan et al. (2021). Although component analyses are integral to our understanding of behavioral interventions, analyses of the components of the CBGG are lacking. Provision of feedback is one component which is easily manipulated and may save teachers time if it is not always necessary. The second aim of this study is therefore to build upon the study conducted by Bohan et al. by evaluating the CBGG-d and CBGG-i across two classroom settings and counterbalancing the intervention conditions. By further demonstrating the efficacy of the CBGG with delayed feedback and reinforcement, we may provide evidence-based game options so that teachers may be granted some autonomy in their own application, without compromising the integrity of the game. Autonomy has emerged in previous research as a core subtheme in influencing teacher “buy-in” or uptake of research practices (Joram et al., 2020) and, as such, is an important consideration when adding to the evidence base. Finally, the study aims to assert whether teachers and students find the CBGG to be an acceptable, effective, and efficient intervention overall.

Research Questions

Research Question 1 (RQ1): What are the effects of the CBGG in targeting academic engagement and disruptive behavior in mainstream Irish adolescent students?

Research Question 2 (RQ2): Is there a difference in effectiveness when the game is played with delayed feedback (CBGG-d) versus immediate feedback (CBGG-i)?

Research Question 3 (RQ3): How acceptable, effective, and efficient is the CBGG according to class teachers? Do class teachers prefer the CBGG-d or the CBGG-i?

Research Question 4 (RQ4): How acceptable is the CBGG according to students? Do students prefer the CBGG-d or the CBGG-i?

Method

Participants and Setting

Participants were two class groups of adolescent students in their first year of secondary school (approximately equivalent to seventh grade in the U.S. school system). Ms. Brady’s class was a mathematics class that consisted of 22 consenting students (10 females, 12 males) with a mean age of 12.7 years. Ms. Brady was a 26-year-old female mathematics teacher with 3 years of experience in teaching. Mr. Carroll’s class was a mathematics class that consisted of 16 consenting students (six females, seven males, and three not reported), with a mean age of 12.8 years. Mr. Carroll was a 23-year-old male mathematics teacher, with 2 years of teaching experience. All students in the respective classes were exposed to the CBGG; however, data were only

collected on consenting students. Neither teacher had implemented the CBGG before; however, Mr. Carroll had witnessed it being implemented while helping out as a team teacher in another classroom in the school. There were no programmed consequences in place for positive behavior in either classroom. Data were collected during regular first-year mathematics instruction that was delivered in line with the Irish curriculum.

Materials

Materials needed for the game were the same for both classrooms. These included laminated copies of the class rules and daily/weekly scoreboards, wearable smart watches (a Fitbit Charge 2 used by Mr. Carroll and an Octopus Watch [Version 1] used by Ms. Brady) that would deliver vibrating prompts, and prizes/reinforcers. Prizes were identified using a preference assessment survey and highly rated prizes included homework passes, free time, sweets, and school cinema passes. The teachers were provided with copies of a procedural checklist. Data were collected using paper and pen, and intervals were signaled through earphones connected to a smartphone.

Dependent Measures

Data were collected on academically engaged behavior (AEB) and disruptive behavior (DB) as dependent measures for this study. AEB was measured across two categories: active engagement and passive engagement, which were considered together for the purposes of graphed data analysis. Active engagement occurred when a student was actively engaged in the ongoing academic task assigned to them, including reading aloud, copying from the board, and talking about the task when this had been permitted. Passive engagement occurred when a student was oriented toward the academic activity but neither actively engaged nor engaging in a defined disruptive behavior. This included behaviors such as looking at the board and reading silently. DB was also considered across two categories that were collated for analysis purposes. These categories were motor and verbal disruption. Motor disruption occurred when a student engaged in a movement not related to the assigned academic task for >3 s consecutively during a 15-s interval. This included being out of seat and playing with objects while not engaged in a task. Verbal disruption occurred when a student engaged in a vocalization not permitted by the teacher and not related to the academic task, including shouting, whispering, and singing.

Observation Procedures and Data Collection

Data were collected up to 5 times per week by the primary observer (first author) during mathematics classes. Data

collection sessions lasted between 15 and 20 min, with the majority of sessions lasting 20 min. The goal for data collection sessions was 20 min; however, on occasion, these were cut short due to competing activities (e.g., the class starting later than planned). The mean observation length in Ms. Brady's class was 18.8 min ($SD = 1.8$ min, range = 15–20 min) and in Mr. Carroll's class was 19.6 min ($SD = 0.8$ min, range = 17–20 min). Data were collected by individual-fixed partial interval recording (DB) and momentary time sampling (AEB). A different student was observed every 15 s in a fixed order around the classroom (Briesch et al., 2015). Data collection took place on consecutive school days, allowing for breaks due to school holidays/planned closures (e.g., Christmas holidays, midterm break).

Design

A multiple treatment reversal design (Cooper et al., 2020), with phases ABACABAC in Ms. Brady's class and ACAB ACABAB in Mr. Carroll's class, was used to assess the effectiveness of the CBGG-d and CBGG-i (A = baseline; B = CBGG-d; C = CBGG-i). Phases were counterbalanced across the classrooms to control for sequence effects. Although an alternating treatments design was considered, the CBGG-d and CBGG-i were deemed too similar, making it difficult for teachers and students to discriminate between the two.

Procedure

Baseline. During the baseline phases, both teachers employed their usual disciplinary procedures while data collection was ongoing. These procedures involved verbal warnings, time out, penalty sheets, and sending students to their form teacher (i.e., the head teacher for the class group), year head (i.e., the head teacher for the year group), or principal.

Teacher Training. The primary researcher (first author) conducted training with both teachers simultaneously during a free class period (approx. 35 min). Game procedures were described with assistance of a PowerPoint presentation and teachers were provided with outlines for implementation in both conditions. The student researcher showed the teachers how to set up and use the wearable prompts. An interval of 5 min was decided as the most reasonable interval length for teachers to conduct behavior checks on teams. Classroom expectations were discussed with both teachers during training and both teachers decided to adopt the same set of expectations, which aligned with target behaviors, for example, "I will respect my classmates and allow them to learn." Points needed to obtain the prizes/reinforcers, which would be available at the end of each game phase, were agreed on by the teacher and researcher and could be adjusted throughout the intervention phases based on

student performance. Teachers were given the opportunity to ask questions at the end of the training session.

Intervention: CBGG. Following a baseline phase, the teachers introduced the CBGG in their classes. There were three teams in Ms. Brady's class and five teams in Mr. Carroll's class. Teams' divisions were decided upon based on the layout of the respective classrooms and particularly disruptive students were dispersed across teams. Apart from the number of teams, procedures for the game were identical across both classrooms.

The class rules and team divisions were explained to students and they were informed that they could earn points for following these class rules. Points could be earned if all members of their team were following the rules when the teacher conducted a "behavior check." These checks occurred, when the teacher was prompted, 6 times during a 40-min class period. The wearable prompting device would vibrate every 5 min and the teacher would check the team's behavior between 0 and 60 s after being prompted, allowing for a variable schedule to be created (e.g., Bohan et al., 2021; Ford, 2017).

Prizes were available weekly after a series of games had been played. The points criterion for the week was set based on how many days the game would be in place and the amount of points that could be possibly earned within that time frame. Points criteria were established in conjunction with the class teacher. Teams meeting the criterion by the end of a series of games earned the top prize for the week. On some occasions, teams had not earned enough points during the week to be in with a chance of winning the weekly prize and, therefore, on occasion, a smaller prize was made available for meeting a smaller daily criterion on the final day of game-play. For example, if the weekly criterion was 35 points and the maximum number of possible points available in a class was 6 points, a team with 28 points at the beginning of Friday's class period could not earn the top prize. In this case, a criterion of 3 points for the class period may be set and a smaller prize would be made available for that game session only. This ensured that teams were still motivated to engage with the game even when the top prize was out of reach. A similar approach was taken by Bohan et al. (2021).

CBGG-d. During the CBGG-d phases, the intervention was implemented as outlined in the previous section and points were recorded discretely by the teachers. Points were made known to the students at the end of each game session and added to the weekly scoreboard.

CBGG-i. During the CBGG-i phases, the intervention was implemented as outlined previously and points were recorded immediately on the daily scoreboard under the team's name. Points were added to the weekly scoreboard at the end of each game session.

Treatment Integrity. The teacher checklist included 11 steps for completion of the game and both the teacher and primary observer had access to this checklist. These steps were as follows: (a) announce that the game will be played, (b) remind the class of team divisions, (c) display the rules, (d) review of the rules by the teacher, (e) remind students how to earn a point, (f) remind students how many points are needed for the weekly prize, (g) announce game start, (h) scan the room when the timer vibrates and award points (privately or publicly dependent on phase), (i) announce the end of the game, (j) announce/tally team points and write them onto the weekly scoreboard, (k) remind students how many points are still needed toward the prize or, during final game day, announce the winners and administer prizes. The primary observer completed the checklist daily during intervention sessions. Treatment integrity data were collected during 100% of intervention sessions in Ms. Brady's class and during 95.65% of sessions in Mr. Carroll's class. The teachers were not required to complete the checklist manually during class, but they were asked to keep it on their desk to refer to during game implementation. If treatment integrity dropped below 80% for more than one session consecutively, this was brought to the teacher's attention in person or through email and they were encouraged to follow all intervention steps on the following days. This type of emailed feedback has been effective in targeting improvements in treatment integrity of primary school teachers implementing the GBG and CBGG (Fallon et al., 2018). Retraining of teachers was not possible due to time constraints during the study. Mean treatment integrity was 79.6% ($SD = 22.1\%$, range = 27.3%–100%) for Ms. Brady and 72.7% for Mr. Carroll ($SD = 18.2\%$, range = 27.3%–100%). Ms. Brady implemented three of the game steps in less than 70% of intervention sessions, namely, reviewing the rules, announcing the game had ended, and reminding students of points needed for the weekly prize. Mr. Carroll implemented five of the game steps in less than 70% of intervention sessions: reviewing the rules, reminding students of the amount of points needed to win, announcing the game had ended, tallying the points on the leader board, and reminding students of points needed for the weekly prize. Evidently, steps missed by both teachers tended to occur toward the end of the game and, anecdotally, the teacher tended to “catch up” on these steps at the beginning of the class the following day (e.g., tallying the points from the day before, before starting a new game).

Data Analysis

The study design was evaluated in line with the What Works Clearinghouse (WWC, 2017, 2020) standards for single-case research. This involves consideration of interobserver agreement (IOA) data, number of phases, and number of data points per phase. These standards have been applied widely in systematic reviews of single-case

research designs to assert design quality of studies (e.g., Bowman-Perrott et al., 2016; Maggin et al., 2012, 2017). Visual analysis of graphed data evaluated level, trend, and consistency of data within and across phases and also the immediacy of effect and rate of overlap between phases.

Effect sizes were calculated using Tarlow's (2017) recommendations for the calculations of Tau, an effect size based on Kendall's Rank Correlation. Tau was calculated for each AB and AC phase contrast using Tarlow's (2016) Baseline-corrected Tau calculator. Weighted mean effect sizes were then calculated for both versions of the game and for both outcome variables by weighting effects for each phase transition by their inverse variances (Tarlow, 2017) and calculating a weighted mean effect size using these weights. For interpretation of Tau, an effect size of .20 may be considered small, .20 to .60 moderate, .60 to .80 large, and .80+ very large (Vannest & Ninci, 2015).

IOA. IOA measures were conducted by trained undergraduate and graduate psychology students during 28.6% of sessions for Ms. Brady's class and 26.7% of sessions for Mr. Carroll's class. IOA was collected at least once per phase and during at least 20% of data points for each condition, as per the WWC design standards (WWC, 2017, 2020). If IOA fell below 80% for any observer, that observer was retrained before they were assigned to collect data again. Mean IOA for AEB was 86.6% ($SD = 4.9\%$, range = 81.3%–95.2%) and 80.9% ($SD = 6.4\%$, range = 69.8%–90%) in Ms. Brady's and Mr. Carroll's classes, respectively. Mean IOA for DB was 90.7% ($SD = 4.1\%$, range = 83.1%–97.5%) and 80.1% ($SD = 8.5\%$, range = 63.5%–92.5%) in Ms. Brady's and Mr. Carroll's classes, respectively. Mean IOA across outcomes, therefore, did not fall below 80%.

Social Validity

Social validity has long been considered essential in application of behavior analytic work, specifically around the behavioral goals, the appropriateness of intervention procedures, and the social importance of the effects (Wolf, 1978). The teachers completed the Behavior Intervention Rating Profile (BIRS; Elliott & Von Brock Treuting, 1991) and students completed a modified version of the Children's Intervention Rating Profile (CIRP; Mitchell et al., 2015; Witt & Elliott, 1985) following the final day of data collection. The BIRS is a rating profile made up of 24 items, all of which are positively phrased (e.g., “This intervention proved effective in helping to change the problem behavior of the classroom”). It incorporates the IRP-15 measure of acceptability (Martens et al., 1985) and adds measures of efficiency and effectiveness. As in previous research (e.g., Ford, 2017), items were modified to reflect application of an intervention to a group rather than an individual child, and to the present/past tense. Items were rated from 1 (*strongly disagree*) to

6 (*strongly agree*). Higher scores indicate more positive perceptions. Elliott and Von Brock Treuting (1991) have established the BIRS as being reliable, with Cronbach's alphas of .97, .92, and .87 for each of the three subscales (acceptability, effectiveness and efficiency, respectively), and an overall Cronbach's alpha coefficient of .97. The modified CIRP included eight items to which students responded "yes" or "no." In keeping with previous research (e.g., Bohan et al., 2021; Mitchell et al., 2015), modifications included changing the tense of the items to past tense, using the term "students" rather than "child," and incorporating an additional item on rewards. Turco and Elliott (1986) reported that the original CIRP had an average coefficient alpha of .86.

Results

Study Design

The design of this study met the WWC standards with reservations (WWC, 2017, 2020). The independent variable was manipulated by the researchers. The collection of IOA data at least 20% of the time overall, and within phases as outlined in the "Method" section, and that the mean IOA

was above 80% across the study meant that the study met the standard requirement relating to IOA. Furthermore, there were three attempts to demonstrate intervention effectiveness. The final standard outlines that there should be five data points per phase in order for the study design to meet the WWC standards fully, or three to four data points per phase to meet the WWC standards with reservations. In this study, across both classrooms, data were collected for a minimum of three data points per phase, and therefore the study meets the WWC standards with reservations.

Student Behavior

Ms. Brady's Mathematics Class. Figure 1 displays behavioral data for Ms. Brady's class.

AEB. During the initial baseline phase, AEB was low in Ms. Brady's mathematics class ($M = 61.2\%$, $SD = 5.5\%$, range = 55%–67.7%). When the CBGG-d was put in place, levels of AEB increased immediately and substantially ($M = 76.6\%$, $SD = 9.5\%$, range = 62.5%–82.4%, +15%¹). Although AEB decreased substantially during Data Point 7 (see Figure 1), this may be attributable to low treatment integrity during this session (54.6%). This was the only data

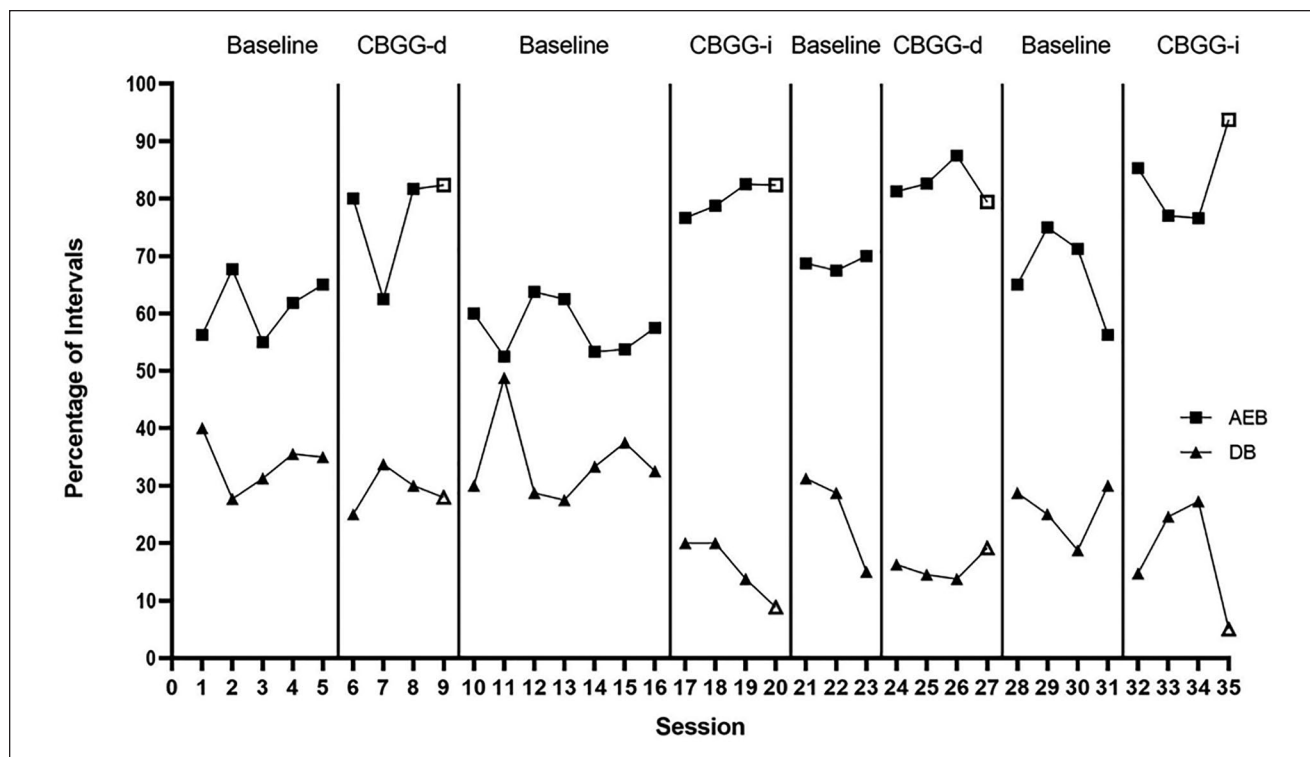


Figure 1. Percentage of Intervals With Academically Engaged and Disruptive Behavior Across Study Phases for Ms. Brady's Mathematics Class.

Note. Open squares and triangles denote sessions where a weekly prize was available (i.e., on a Friday); AEB = academically engaged behavior; DB = disruptive behavior; CBGG-d = intervention Condition 1, consisting of teacher's private recording of points and feedback on team progress toward a weekly goal, which is withheld until the end of the game period; CBGG-i = intervention Condition 2, consisting of teacher's recording of points publicly during the game.

Table 1. Tau Effect Sizes for AEB and DB Across Phase Changes in Ms. Brady's and Mr. Carroll's Mathematics Classes.

| Phase change | Ms. Brady's class | | Phase change | Mr. Carroll's class | |
|------------------------|-------------------|-------|------------------------|---------------------|--------|
| | AEB | DB | | AEB | DB |
| Baseline to CBGG-d | .596 | -.447 | Baseline to CBGG-i | .646 | -.439 |
| Withdrawal to CBGG-i | .714* | -.72* | Withdrawal to CBGG-d | .756* | -.784* |
| Withdrawal 2 to CBGG-d | .756 | -.504 | Withdrawal 2 to CBGG-i | .701* | -.582* |
| Withdrawal 3 to CBGG-i | .756* | -.472 | Withdrawal 3 to CBGG-d | .529 | -.567 |
| — | | | Withdrawal 4 to CBGG-d | .643* | -.447 |
| Weighted mean CBGG-d | .68 | -.47 | Weighted mean CBGG-d | .66 | -.64 |
| Weighted mean CBGG-i | .73 | -.64 | Weighted mean CBGG-i | .68 | -.53 |

Note. AEB = academically engaged behavior; DB = disruptive behavior; CBGG = Caught Being Good Game; CBGG-d = intervention Condition 1, consisting of teacher private recording of points and feedback on team progress toward a weekly goal, which is withheld until the end of the game period; CBGG-i = intervention Condition 2, consisting of teacher's recording of points publicly during the game.

* $p < .05$.

point to overlap with the preceding baseline phase.

During the withdrawal phase, AEB decreased immediately and substantially. It remained relatively stable throughout this phase ($M = 57.6\%$, $SD = 4.6\%$, range = 52.5%–63.8%, –22.4%). There was a large increase in AEB when the CBGG-i was put in place ($M = 80.1\%$, $SD = 2.9\%$, range = 76.7%–82.5%, +19.2%). Furthermore, there was no overlap between this phase and the previous phase.

In the second withdrawal phase, AEB decreased immediately and remained stable across the phase ($M = 68.8\%$, $SD = 1.3\%$, range = 67.5%–70%, –13.6%). The CBGG-d was reinstated following this withdrawal phase and AEB increased immediately, remaining high and stable across the phase ($M = 82.7\%$, $SD = 3.5\%$, range = 79.4%–87.5%, +11.3%). No overlap was observed between this intervention phase and the withdrawal phase that preceded it.

During the final withdrawal phase, there was an immediate decrease in AEB; however, data were not as stable as in previous withdrawal phases ($M = 66.9\%$, $SD = 8.2\%$, range = 56.3%–75%, –14.4%). There was no overlap between this withdrawal phase and the preceding CBGG-d phase. The CBGG-i was implemented in the final phase. There was an immediate and substantial increase in AEB ($M = 83.2\%$, $SD = 8.1\%$, range = 76.6%–93.8%, +29%), and there was no overlap with the preceding withdrawal phase.

Overall, there were consistent changes in behavior when either version of the CBGG was in place. The magnitude of behavior change across conditions could be considered medium, characterized by immediate and generally stable increases in AEB in intervention phases compared with baseline phases.

Disruptive Behavior. DB was high during the initial baseline phase, occurring during a mean of 33.9% of intervals ($SD = 4.7\%$, range = 27.7%–40%). DB decreased immediately when the CBGG-d was introduced; however,

data did not remain low and stable and the overall change in level was minimal ($M = 29.2\%$, $SD = 3.7\%$, range = 25%–33.8%, –10%). Three of the four data points overlapped with the data in the baseline phase.

When the CBGG-d was withdrawn, changes in DB were not very pronounced and remained at a level similar to the previous phase ($M = 34.1\%$, $SD = 7.3\%$, range = 27.5%–48.8%, +2.1%). Upon introduction of the CBGG-i for the first time, there was an immediate and substantial decrease in DB ($M = 15.6\%$, $SD = 5.4\%$, range = 8.8%–20%, –12.5%), with a decreasing trend across the phase. There was no overlap with the preceding withdrawal phase.

The CBGG-i was subsequently withdrawn and DB increased immediately; however, it decreased again toward the end of the phase, with a general downward trend across the phase ($M = 25\%$, $SD = 8.8\%$, range = 15%–31.3%, +22.4%). When the CBGG-d was reinstated, DB remained low and stable ($M = 15.9\%$, $SD = 2.4\%$, range = 13.8%–19.1%, +1.3%); however, 50% of the data points overlapped with the preceding withdrawal phase.

DB increased modestly to a mean of 25.6% of intervals ($SD = 5.1\%$, range = 18.8%–30%, +9.6%) during the following, final withdrawal phase. The CBGG-i was implemented during the final phase and DB decreased immediately ($M = 17.9\%$, $SD = 10.2\%$, range = 5%–27.3%, –15.3%). However, across the phase, two data points saw DB increase to levels similar to the previous withdrawal phase, which resulted in 50% overlap.

Overall, the magnitude of change in DB could be considered small. Immediate changes in behavior were not present during every phase change, despite general changes in level of DB being present.

Effect Size Data. Tau effect sizes across phase changes in Ms. Brady's class are presented in Table 1. Weighted average Tau effect sizes for the impact of the CBGG-d and CBGG-i on AEB were .68 and .73, respectively. These are

considered large effect sizes. Weighted average Tau effect sizes for the impact of the CBGG-d and CBGG-i on DB were $-.47$ (moderate) and $-.64$ (large), respectively. Overall, the CBGG-i produced larger effect sizes for both target behaviors across phase changes.

Mr. Carroll's Mathematics Class. Figure 2 is a graph depicting AEB and DB across phases in Mr. Carroll's mathematics class.

AEB. AEB in Mr. Carroll's class was relatively stable at baseline, occurring during a mean of 57.8% of intervals ($SD = 6.2\%$, range = 52.5%–67.5%). When the CBGG-i was introduced, there was an immediate increase in AEB ($M = 73.9\%$, $SD = 5.9\%$, range = 67.1%–77.5%, +24.7%). Despite this immediate increase during the first two data points in this phase, there was a subsequent decrease during the third data point and this data point overlapped with the baseline phase.

Despite a decrease in AEB toward the end of the previous intervention phase, there was an even larger decrease in AEB at the beginning of the first withdrawal

phase ($M = 46.6\%$, $SD = 12.5\%$, range = 31.4%–57.5%, -25.8%). The CBGG-d was then put in place for the first time and there was an increase in AEB ($M = 70.3\%$, $SD = 3.7\%$, range = 66.3%–75%, +10%), with no data points overlapping with the preceding withdrawal phase. The data remained stable across this phase.

When the CBGG-d was withdrawn, AEB immediately decreased to levels similar to the initial baseline phase ($M = 54.5\%$, $SD = 2.9\%$, range = 51.3%–57.5%, -11.3%) and remained stable. The CBGG-i was implemented for a second time in the following phase. Mr. Carroll mistakenly implemented the CBGG-d during Data Point 22 (see Figure 2). To ensure that enough data were collected on the CBGG-i during the phase (i.e., >3 points), the phase was extended across 2 weeks. There was a change in level in AEB ($M = 69.4\%$, $SD = 8.5\%$, range = 57.5%–81.3%, +26%); however, data were highly variable across this phase, with a general downward trend for AEB. Despite this trend and highly variable data, only one data point (Data Point 26; see Figure 2) overlapped with the previous phase, meaning there was a high percentage of nonoverlapping data (87.5%).

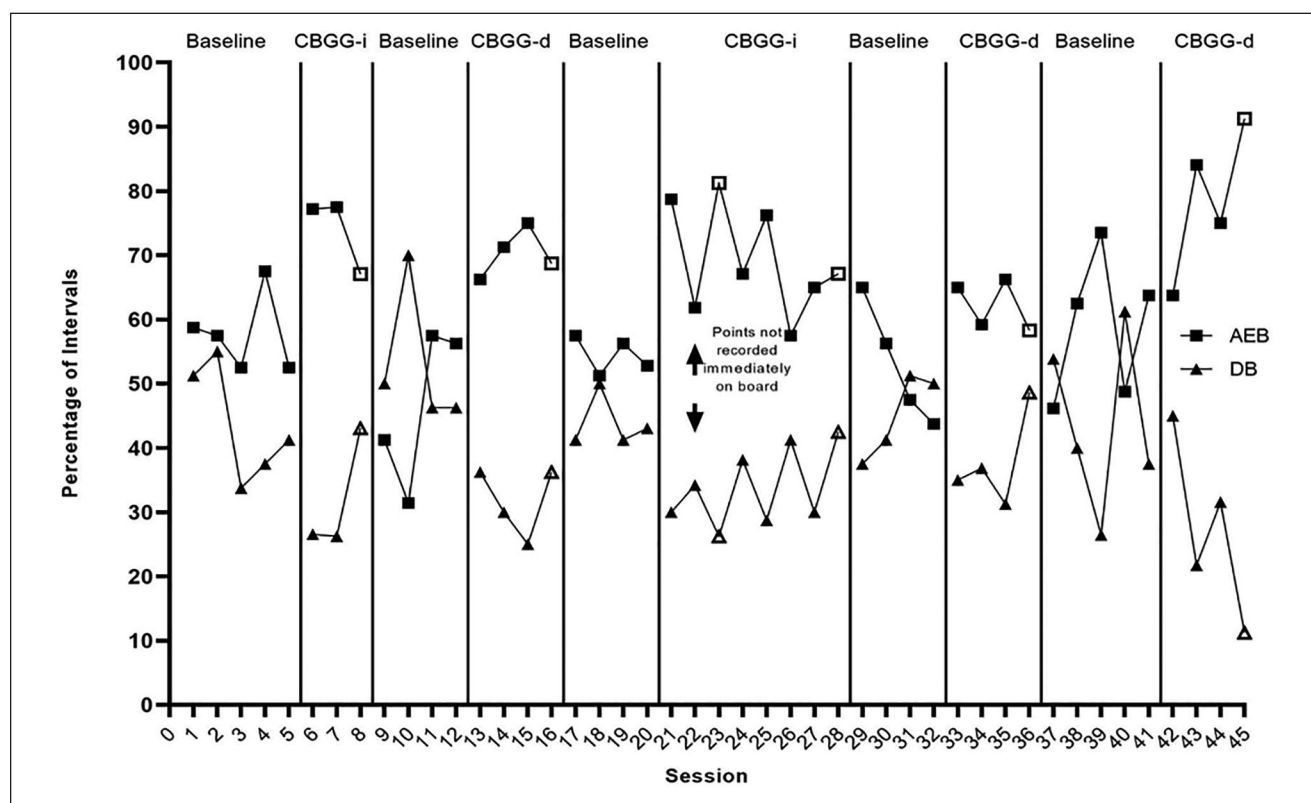


Figure 2. Percentage of Intervals With Academically Engaged and Disruptive Behavior Across Study Phases for Mr. Carroll's Mathematics Class.

Note. Open squares and triangles denote sessions where a weekly prize was available (i.e., on a Friday); AEB = academically engaged behavior; DB = disruptive behavior; CBGG-d = intervention Condition 1, consisting of teacher's private recording of points and feedback on team progress toward a weekly goal, which is withheld until the end of the game period; CBGG-i = intervention Condition 2, consisting of teacher's recording of points publicly during the game.

During the next withdrawal phase, AEB continued to decrease at a much steeper rate than during the previous CBGG-i phase ($M = 53.1\%$, $SD = 9.5\%$, range = 43.8%–65%, -2.1%). There was an immediate improvement evident in AEB when the CBGG-d was put in place for the second time ($M = 62.2\%$, $SD = 4\%$, range = 58.3%–66.3%, $+21.3\%$) and this remained stable across the phase. However, there was a high percentage of overlap between this phase and the preceding withdrawal phase (75%). During this phase, teams performed very poorly in earning points for the game. It was therefore decided that the game would be extended into the following week with a different prize. However, Mr. Carroll forgot to implement the game on the first day of the following week. For this reason, an additional baseline phase was put in place, before trialing the CBGG-d for a final time.

Behavior was highly variable during the final withdrawal phase. AEB occurred during a mean of 58.9% of intervals ($SD = 11.4\%$, range = 46.2%–73.5%, -12.2%). During the third data point in this phase (Point 39; see Figure 2), there were two team teachers present (i.e., additional teachers working in a support capacity with Mr. Carroll). This was a confounding factor not apparent during any previous data collection session, which may have unduly impacted upon AEB. In the final intervention phase where the CBGG-d was put in place, there was a steady increase in AEB across the week ($M = 78.5\%$, $SD = 11.9\%$, range = 63.8%–91.3%, $+0\%$). However, changes were not immediate and there was a degree of overlap with the previous phase. During this week, students were approaching a school break and absenteeism was high. As well as this, two team teachers were present again during Data Points 43 and 45 (see Figure 2). Therefore, only tentative conclusions must be drawn for this phase change due to the confounding variables.

Overall, small to medium increases in AEB were observed during each phase change from baseline/withdrawal to intervention phases. The magnitude of behavior change may therefore be characterized as small to medium.

Disruptive Behavior. During the initial baseline phase, DB was variable; however, it demonstrated an upward trend toward the end of the phase, across the final three data points ($M = 43.8\%$, $SD = 9.1\%$, range = 33.8%–55%). The CBGG-i was introduced and DB decreased to a mean of 32% of intervals ($SD = 9.6\%$, range = 26.3%–43%, -14.7%). The data did not remain stable and the final data point in this phase saw an increase in DB, which overlapped with the baseline phase.

The CBGG-i was withdrawn and there was an increase in DB ($M = 53.1\%$, $SD = 11.4\%$, range = 46.3%–70%, $+7\%$). There was no overlap here with the preceding intervention phase. The CBGG-d was introduced during the following phase. DB decreased immediately ($M = 31.9\%$,

$SD = 5.5\%$, range = 25%–36.3%, -10%). DB remained moderately stable and did not overlap with the previous withdrawal phase.

When the CBGG-d was withdrawn, DB increased to a stable level, and did not overlap with the preceding intervention phase ($M = 43.9\%$, $SD = 4.2\%$, range = 41.3%–50%, $+5\%$). The CBGG-i was implemented in the following phase and, as previously mentioned, Mr. Carroll implemented the CBGG-d during Data Point 22 in error (see Figure 2). DB was low initially in this phase; however, it increased across the phase and remained unstable ($M = 33.9\%$, $SD = 6.1\%$, range = 26.3%–42.5%, -13.1%). Data points 26 and 28 overlapped with the preceding withdrawal phase, meaning the percentage of nonoverlapping data was 75%.

Although DB did not increase immediately during the following withdrawal phase, behavior increased steadily across the phase ($M = 45\%$, $SD = 6.7\%$, range = 37.5%–51.3%, -5%). The CBGG-d was then put in place and an immediate decrease in DB was evident ($M = 37.9\%$, $SD = 7.5\%$, range = 31.3%–48.6%, -15%). However, during the final data point, there was an increase to a level similar to high levels during baseline phases.

DB occurred during a mean of 43.8% of intervals ($SD = 13.8\%$, range = 26.5%–61.3%, $+5.2\%$) during the final withdrawal phase. The data were unstable here with extreme highs and lows evident in DB. There was a decrease in DB, albeit not immediate, when the CBGG-d was put in place ($M = 27.4\%$, $SD = 14.4\%$, range = 11.3%–45%, $+7.5\%$); however, this must be considered in light of the confounds reported in the “Academically Engaged Behavior” section earlier.

Immediate decreases in DB were observed for five phase changes from baseline to intervention; however, some of these decreases were small and data were unstable at times. Therefore, the magnitude of behavior change may be classified as being small.

Effect Size Data. Tau effect sizes across phase changes and weighted mean Tau effect sizes for each version of the game are presented in Table 1. Data Point 22 was omitted from calculations due to the CBGG-d being implemented during this data point. The data point therefore did not align with the other data points in the phase. Weighted average Tau effect sizes for the impact of the CBGG-d and the CBGG-i on AEB in Mr. Carroll’s class were .66 and .68, respectively. These are considered large effect sizes. Weighted average Tau effect sizes for the impact of the CBGG-d and CBGG-i on DB were $-.64$ and $-.53$, respectively. The effect size was therefore large for the CBGG-d and moderate for the CBGG-i. The CBGG-d and CBGG-i therefore had similar effects on AEB, whereas the CBGG-d had a slightly larger effect than the CBGG-i on DB in this class group.

Social Validity

Teacher Social Validity. Ms. Brady responded predominantly negatively in completing the BIRS (Elliott & Von Brock Treuting, 1991), with a mean rating of 2.5 for acceptability, 1 for effectiveness, and 1 for efficiency. In her written feedback, she stated that *“there were too many steps involved and it adds an extra layer of work to an already over-stretched teacher.”* She also noted concerns over student behavior when observers were present in the room, the game’s fairness, and did not think it should be referred to as “a game.” Despite negative feedback, she also referenced the fact that students *“. . . are constantly asking to play the game,”* which may indicate student approval. She preferred the CBGG-d over the CBGG-i, stating that *“publicly showing points caused [students] to ask why they didn’t get one.”*

Mr. Carroll rated the game much more positively. He rated the intervention with a mean of 5.1 for acceptability, 4 for effectiveness, and 5 for efficiency. In his written feedback, he stated that he thought the game was *“definitely a worthwhile intervention”* and that *“the students did react positively.”* He preferred the CBGG-d, stating that it *“means [students] can’t question the decision and they’re less likely to give up on the game once the teacher keeps reminding them of the game.”*

Student Social Validity. Nineteen students in Ms. Brady’s class and 14 students in Mr. Carroll’s class completed the modified CIRP (Mitchell et al., 2015; Witt & Elliott, 1985). Students in Ms. Brady’s class rated the CBGG moderately positively on the CIRP ($M = 4.8$). A majority of students responded positively to at least six of the eight statements. On the items where a majority did not respond positively, 12 students disagreed that the game helped them to do better in mathematics class and 10 students thought the game may have caused problems for their classmates. A total of 42.1% of the students preferred the CBGG-i and 42.1% preferred the CBGG-d. Two students did not report a preference. Mr. Carroll’s class rated the game more positively than Ms. Brady’s class, scoring a mean of 6.4. A majority of students rated each item positively rather than negatively. The largest proportion of negative responses was for the item “Do you think the game caused any problems for your classmates?” to which six students responded “yes.” None of the students elaborated on potential problems that could have been caused by the game. A total of 64.3% of the students preferred the CBGG-i and 28.6% preferred the CBGG-d. One student did not report a preference.

Discussion

This study aimed to build upon the work of Bohan et al. (2021) by evaluating the CBGG-d and the CBGG-i across

two secondary school classrooms, while counterbalancing presentation of intervention conditions, using a multiple treatment reversal design. Across both classrooms, improvements were evident in AEB and DB when the CBGG was in place. There were differing effects apparent for the CBGG-d and CBGG-i across classrooms, with the CBGG-i appearing slightly more effective in Ms. Brady’s classroom and the CBGG-d appearing more effective in Mr. Carroll’s classroom.

In Ms. Brady’s class, AEB was almost always higher in phases where the CBGG was in place compared with baseline and withdrawal phases. This is also reflected in the large Tau effect sizes for the impact of the CBGG-d and the CBGG-i on AEB (.68 and .73, respectively). Differences between the two versions of the game were minimal, with the CBGG-i appearing very slightly more effective in targeting AEB. In Ms. Brady’s class, mean rates of DB were always lower in intervention phases when compared with preceding baseline/withdrawal phases. Initial changes in DB were minimal. The CBGG-d introduced in the first intervention phase did not produce a substantial and immediate change in behavior when compared with baseline, and this is reflected in a moderate effect size for the phase change ($\text{Tau} = -.447$). As the study progressed, moderate to large changes in DB were evident between phases and reductions were evident when both versions of the game were put in place. Despite this, a downward trend was observed during the second withdrawal phase, which saw the phase end with relatively low DB. Interestingly, there was no corresponding increase in AEB. Changes in DB were not as stable or substantial in this class group when compared with changes in AEB. This is reflected in slightly smaller effect sizes for the CBGG-d (weighted mean $\text{Tau} = -.47$) and the CBGG-i (weighted mean $\text{Tau} = -.64$). The CBGG-i again appeared slightly more effective than the CBGG-d in targeting DB in Ms. Brady’s class.

In Mr. Carroll’s class, mean rates of AEB during intervention phases always exceeded mean rates during baseline/withdrawal phases. There were data points during baseline/withdrawal phases; however, where AEB was quite high (e.g., Data Point 39; see Figure 2), meaning there was a high rate of overlap overall between data points in withdrawal phases and intervention phases. Overall, large effect sizes were identified for the impact of the CBGG-d (weighted mean $\text{Tau} = .66$) and CBGG-i (weighted mean $\text{Tau} = .68$) on AEB, and differences between the two versions of the game were minimal. Effects of the CBGG on DB were mainly positive. There was always a decrease in mean DB when a game phase was introduced versus the preceding baseline or withdrawal phase. There were some phases, however, where upward trends in DB were observed despite the CBGG being in place (e.g., second implementation of the CBGG-i). Despite an initial decrease in the following phase, this upward trend continued,

suggesting that perhaps there were other environmental variables at play contributing to behavior. Toward the end of the study, behavior became very unstable, particularly in the final two phases. Very low rates of DB in these phases are potentially attributable to the presence of extra teachers in the room as part of a team-teaching program employed in the school. Overall, moderate to large effect sizes were observed across phase changes, and weighted mean effect sizes for each version of the game were $-.64$ (CBGG-d) and $-.53$ (CBGG-i). Differences between the two versions of the game were small; however, the CBGG-d was slightly more effective in targeting DB. This finding must be considered in light of the confounding factors mentioned in the “Results” section; two additional team teachers were present during some CBGG-d phases toward the end of the study, which may have influenced student DB.

Taken together, the results align with the findings of Bohan et al. (2021) and with previous research demonstrating the efficacy of group contingencies in general (Maggin et al., 2012, 2017), and the CBGG specifically, in targeting classroom-based behavior (e.g., Ford, 2017; Tanol et al., 2010; Wahl et al., 2016; Wright & McCurdy, 2012). The findings lend further support to the efficacy of the CBGG with delayed feedback (e.g., Wahl et al., 2016) and immediate visual feedback (e.g., Lynne et al., 2017). Bohan et al. (2021) employed a reversal design in one classroom, which did not allow for counterbalancing of intervention conditions and therefore could not control for sequence effects. This study controlled for this, thereby advancing on previous research and demonstrating that it was not necessarily the first version of the CBGG that was implemented and which was the most effective across classrooms. In fact, the contrary was true in Ms. Brady’s classroom where the CBGG-d was implemented first, but the CBGG-i appeared slightly more effective overall. The CBGG-i was implemented first in Mr. Carroll’s classroom, yet the CBGG-d, at times, appeared slightly more effective. Taken together, these findings suggest that sequence effects were not present when implementing the CBGG with or without feedback, and other variables may be at play which influence the relative effectiveness of either version with particular class groups (e.g., students’ reinforcement history). Future research may consider which other contextual variables may influence the effectiveness of the CBGG-d and the CBGG-i. Another point of note is that, in general, across both classrooms, both versions of the CBGG appeared to have a larger impact on AEB than on DB. Previous research has identified patterns whereby the GBG has had a larger effect on DB than AEB (e.g., Ford et al., 2020). The opposite was evident here, suggesting that, although students tended to be more engaged during the CBGG, periods of disruption still did occur. Data on DB was collected using partial interval recording, meaning that it was coded as occurring any time within a 15-s interval.

AEB on the contrary, was documented using momentary time sampling. It is therefore possible that a student could be engaging with DB early during a 15-s interval, but may have reverted to AEB by the end of the interval.

Social Validity of the CBGG With Teachers and Students

Ms. Brady rated the CBGG predominantly negatively, citing time constraints as a critical issue when giving reasons for her unfavorable rating. This negative rating aligns with recent research conducted on the GBG with primary school students (Dadakhodjaeva et al., 2020). Mr. Carroll rated the game very positively and called it a “*worthwhile intervention*” for his class. This rating compares favorably with recent positive ratings of the CBGG (Bohan et al., 2021) and the GBG (Mitchell et al., 2015) with adolescent students, and potentially reflects Mr. Carroll’s satisfaction with the CBGG’s impact on behavior in his classroom. Ms. Brady’s rating and subsequent comments can be considered in conjunction with recent research carried out in an Irish context citing that Irish second-level teachers are overstretched and under time pressure during the school day (The Association of Secondary Teachers, Ireland (ASTI), 2018). Perhaps a less time-intensive intervention or a convention whereby the game is played at a reduced frequency (as demonstrated with the GBG by Dadakhodjaeva et al., 2020) would be more desirable for Ms. Brady. It is also possible that Ms. Brady paid more attention to acute disruptive behaviors of individual students, or behavior of students not consenting to participating in the research (which would not have been captured in the data) than to more positive class-wide shifts in behavior demonstrated in the data collected. Both Ms. Brady and Mr. Carroll preferred the CBGG-d to the CBGG-i, which is in keeping with teacher preferences reported previously (Bohan et al., 2021), and is likely linked to the fact that it involves less distraction for the teacher and students. Students in both classes rated the CBGG positively, with Mr. Carroll’s class rating the game slightly more positively than Ms. Brady’s students. Interestingly, this corresponds with the classes’ respective teachers’ ratings. The positive ratings align with much previous research in this area, in that students have rated the CBGG favorably (e.g., Wahl et al., 2016; Wright & McCurdy, 2012). Future research may consider collecting social validity data intermittently throughout a study rather than solely at the end of the study to give teachers an opportunity to raise concerns.

Implications for Practice

These findings lend further support to the efficacy of the CBGG with adolescent students and, importantly, have

provided further insight into a component that is potentially variable, that is, feedback. The findings provide further support for a positive classroom management intervention that consistently draws teacher attention to positive and appropriate behavior. The similar effectiveness of the CBGG-d and CBGG-i points toward the potential for teachers to be flexible in their approach to using the CBGG. This may be useful for teachers who are under time pressure and do not have time in their schedules to engage with more intensive behavior support plans and individualized classroom management. Although the CBGG has been evaluated with adolescent students before, previous versions have incorporated ClassDojo technology (e.g., Dadakhodjaeva, 2017; Ford, 2017). The version applied here used a simpler, low-tech method of awarding points, which may be more desirable for use by teachers in secondary school classrooms who have limited time frames in which to carry out class duties. The teachers in this study preferred the CBGG-d to the CBGG-i, so it is possible that they would have preferred this version over a version incorporating ClassDojo. These findings also have implications for flexibility and teacher autonomy in intervention implementation. Taken together with the findings from Bohan et al. (2021), teachers at the lower secondary school level may consider implementing the CBGG with or without immediate visual feedback, depending on their own preferences or perhaps the preferences of their class. This fostering of autonomy may have important implications for enhancing teacher uptake of research-based practices (Joram et al., 2020). Of importance to note, is the Irish context in which this study was implemented. This was only the second study applying the CBGG with Irish students (the first application being that by Bohan et al. 2021). First-year students in the Irish school context have recently undergone a period of transition from 8 years of primary school to secondary, which encompasses many structural changes, for example, a move from having one teacher all day to moving from classroom to classroom with multiple teachers in one day. Given this transition, and the potential challenges faced during this time, first-year class groups are suitable targets for classroom management interventions such as the CBGG.

Study Limitations and Future Directions

These findings must be evaluated in light of some limitations. First, the results are relevant to a first-year mathematics classroom setting, such that both classes in this study maintained these characteristics. Future research may consider the effectiveness of the CBGG with and without feedback with other classes in a secondary school, for example, older adolescent students. Second, only visual feedback was considered in this study and therefore conclusions cannot be drawn about the relative importance

of vocal feedback. Wiskow et al. (2019) evaluated the GBG with conditions including no feedback, visual feedback only, vocal feedback only, and visual + vocal feedback. Replication of this study with the CBGG and older students may be a useful avenue for future research. A third limitation is that the study does not meet the WWC standards (2017, 2020) fully; however, it does meet the WWC design standards with reservations. This designation means that the study maintained reasonably high quality in design standards, with some issues. Data were collected during three to four data points per phase on occasion throughout the study; however, this reflects the difficulties in collecting data in a classroom setting. A strength of single-case research designs is their applicability in applied settings (Kazdin, 2011); however, consideration must be given to the time constraints within these settings when evaluating study designs. Finally, treatment integrity was low on several occasions throughout this study. It was apparent that the final steps of the CBGG were often missed by teachers due to time constraints at the end of class. A feedback mechanism was in place for teachers during this study (emailed or in person feedback); however, a more potent method may be required in future (e.g., prompts to complete steps of the game in real time).

Conclusion

These findings provide further support for the effectiveness of the CBGG generally with adolescent students. This classroom management intervention with a positive focus may be useful for teachers wishing to adopt such approaches in their classrooms, for example, in schools adopting a school-wide positive behavior approach. Furthermore, it has provided further insight into the role of visual feedback during the game, suggesting that it is not always a necessary component. These findings are situated in a growing body of research on the efficacy of versions of the GBG maintaining a focus on positive reinforcement only, rather than a combination of reinforcement and punishment. The findings importantly also point toward scope for teacher autonomy in implementation, such that teachers may choose to implement the CBGG with or without feedback, or with weekly, rather than, daily prizes. Future research may consider expanding on this study by recruiting more diverse secondary school classes (e.g., older adolescents, subjects other than mathematics) and evaluating feedback types such as vocal feedback or visual + vocal feedback.

Authors' Note

This research was completed as part of the first author's PhD thesis research. This study was conducted in accordance with full ethical approval granted by Dublin City University's Research Ethics Committee.

Acknowledgments

The authors would like to acknowledge the assistance of DCU psychology students for their help in collecting data for this project.

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 was partially funded by a Dublin City University Career Enhancement Grant awarded to senior author and primary research supervisor, Dr Sinéad Smyth.

Note

1. This refers to the percentage change from the final data point in the previous phase to the first data point in the phase under discussion.

References

- Association of Secondary Teachers, Ireland. (2018). *Teachers' Work: Work Demands and Work Intensity, March 2018*. <https://www.asti.ie/document-library/teachers-work-work-demands-and-work-intensity-march-2018/#:~:text=89%25%20of%20teachers%20agree%20that,activities%20during%20the%20school%20day.&text=97%25%20of%20teachers%20say%20their,2009%20to%2051%25%20in%202018>.
- Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good Behavior Game: Effect of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis*, 2(2), 119–124. <https://doi.org/10.1901/jaba.1969.2-119>
- Bohan, C., Smyth, S., & McDowell, C. (2021). An evaluation of the Caught Being Good Game with an adolescent student population. *Journal of Positive Behavior Interventions*, 23(1), 42–52. <https://doi.org/10.1177/1098300720928455>
- Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., & Vannest, K. (2016). Promoting positive behavior using the Good Behavior Game: A meta-analysis of single-case research. *Journal of Positive Behavior Interventions*, 18(3), 180–190. <https://doi.org/10.1177/1098300715592355>
- Briesch, A. M., Hemphill, E. M., Volpe, R. J., & Daniels, B. (2015). An evaluation of observational methods for measuring response to classwide intervention. *School Psychology Quarterly*, 30(1), 37–49. <https://doi.org/10.1037/spq0000065>
- Cooper, J. O., Heward, W. L., & Heron, T. E. (2020). *Applied behavior analysis* (3rd ed.). Pearson Education.
- Dadakhodjaeva, K. (2017). *The Good Behavior Game: Effects on and maintenance of behavior in middle-school classrooms using Class Dojo* [Doctoral dissertation]. University of Southern Mississippi. <https://aquila.usm.edu/dissertations/363/>
- Dadakhodjaeva, K., Radley, K. C., Tingstrom, D. H., Dufrene, B. A., & Dart, E. H. (2020). Effects of daily and reduced frequency implementation of the Good Behavior Game in kindergarten classrooms. *Behavior Modification*, 44(4), 471–495. <https://doi.org/10.1177/0145445519826528>
- Donaldson, J., Fisher, A. B., & Kahng, S. (2017). Effects of the Good Behavior Game on individual student behavior. *Behavior Analysis: Research and Practice*, 17(3), 207–216. <https://doi.org/10.1037/bar0000016>
- Elliott, S. N., & Von Brock Treuting, M. (1991). The Behavior Intervention Rating Scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology*, 29(1), 43–51. [https://doi.org/10.1016/0022-4405\(91\)90014-I](https://doi.org/10.1016/0022-4405(91)90014-I)
- Fallon, L. M., Collier-Meek, M. A., Kurtz, K. D., & DeFouw, E. R. (2018). Emailed implementation supports to promote treatment integrity: Comparing the effectiveness and acceptability of prompts and performance feedback. *Journal of School Psychology*, 68, 113–128. <https://doi.org/10.1016/j.jsp.2018.03.001>
- Foley, E. A., Dozier, C. L., & Lessor, A. L. (2019). Comparison of components of the Good Behavior Game in a preschool classroom. *Journal of Applied Behavior Analysis*, 52(1), 84–104. <https://doi.org/10.1002/jaba.506>
- Ford, W. B. (2017). *Evaluation of a positive version of the Good Behavior Game utilizing ClassDojo technology in secondary classrooms* [Doctoral dissertation]. University of Southern Mississippi. <http://aquila.usm.edu/dissertations>
- Ford, W. B., Radley, K. C., Tingstrom, D. H., & Dufrene, B. A. (2020). Efficacy of a no-team version of the Good Behavior Game in high school classrooms. *Journal of Positive Behavior Interventions*, 22(3), 181–190. <https://doi.org/10.1177/1098300719890059>
- Groves, E. A., & Austin, J. L. (2020). Examining adaptations of the Good Behavior Game: A comparison of known and unknown criteria for winning. *School Psychology Review*, 49(1), 74–84. <https://doi.org/10.1080/2372966X.2020.1716635>
- Hernan, C. J., Collins, T. A., Morrison, J. Q., & Kroeger, S. D. (2019). Decreasing inappropriate use of mobile devices in urban high school classrooms: Comparing an antecedent intervention with and without the Good Behavior Game. *Behavior Modification*, 43(3), 439–463. <https://doi.org/10.1177/0145445518764343>
- Joram, E., Gabriele, A. J., & Walton, K. (2020). What influences teachers' "buy-in" of research? Teachers' beliefs about the applicability of educational research to their practice. *Teaching and Teacher Education*, 88, 1–12. <https://doi.org/10.1016/j.tate.2019.102980>
- Joslyn, P. R., Donaldson, J. M., Austin, J. L., & Vollmer, T. R. (2019). The Good Behavior Game: A brief review. *Journal of Applied Behavior Analysis*, 52(3), 811–815. <https://doi.org/10.1002/jaba.572>
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). Oxford University Press.
- Kleinman, K. E., & Saigh, P. A. (2011). The effects of the Good Behavior Game on the conduct of regular education New York City high school students. *Behavior Modification*, 35(1), 95–105. <https://doi.org/10.1177/0145445510392213>
- Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied*

- Behavior Analysis*, 8(3), 341–347. <https://doi.org/10.1901/jaba.1975.8-341>
- Lynne, S., Radley, K. C., Dart, E. H., Tingstrom, D. H., Barry, C. T., & Lum, J. D. (2017). Use of a technology-enhanced version of the Good Behavior Game in an elementary school setting. *Psychology in the Schools*, 54(9), 1049–1063. <https://doi.org/10.1002/pits.22043>
- 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(5), 625–654. <https://doi.org/10.1016/j.jsp.2012.06.001>
- 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. K., Witt, J. C., Elliott, 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>
- McKenna, J. W., & Flower, A. (2014). Get them back on track: Use of the Good Behavior Game to improve student behavior. *Beyond Behavior*, 23(2), 20–26. <https://doi.org/10.1177/107429561402300204>
- Mitchell, R. R., Tingstrom, D. H., Dufrene, B. A., Ford, W. B., & Sterling, H. E. (2015). The effects of the Good Behavior Game with general-education high school students. *School Psychology Review*, 44(2), 191–207. <https://doi.org/10.17105/spr-14-0063.1>
- Nolan, J. D., Filter, K. J., & Houlihan, D. (2014). Preliminary report: An application of the Good Behavior Game in the developing nation of Belize. *School Psychology International*, 35(4), 421–428. <https://doi.org/10.1177/0143034313498958>
- Stromer, R., McComas, J. J., & Rehfeldt, R. A. (2000). Designing interventions that include delayed reinforcement: Implications of recent laboratory research. *Journal of Applied Behavior Analysis*, 33(3), 359–371. <https://doi.org/10.1901/jaba.2000.33-359>
- Tanol, G., Johnson, L., McComas, J., & Cote, E. (2010). Responding to rule violations or rule following: A comparison of two versions of the Good Behavior Game with kindergarten students. *Journal of School Psychology*, 48(5), 337–355. <https://doi.org/10.1016/j.jsp.2010.06.001>
- Tarlow, K. R. (2016). *Baseline corrected tau calculator*. <http://www.ktarlow.com/stats/tau>
- Tarlow, K. R. (2017). An improved rank correlation effect size statistic for single-case designs: Baseline corrected Tau. *Behavior Modification*, 41(4), 427–467. <https://doi.org/10.1177/0145445516676750>
- Turco, T. L., & Elliott, S. N. (1986). Students' acceptability ratings of interventions for classroom misbehaviors: A study of well-behaving and misbehaving youth. *Journal of Psychoeducational Assessment*, 4(4), 281–289. <https://doi.org/10.1177/073428298600400404>
- Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling and Development*, 93(4), 403–411. <https://doi.org/10.1002/jcad.12038>
- Wahl, E., Hawkins, R. O., Haydon, T., Marsicano, R., & Morrison, J. Q. (2016). Comparing versions of the Good Behavior Game: Can a positive spin enhance effectiveness? *Behavior Modification*, 40(4), 493–517. <https://doi.org/10.1177/0145445516644220>
- What Works Clearinghouse. (2017). *What Works Clearinghouse™ standards handbook* (Version 4.0). https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf
- What Works Clearinghouse. (2020). *What Works Clearinghouse™ standards handbook* (Version 4.1). <https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf>
- Wiskow, K. M., Matter, A. L., & Donaldson, J. M. (2019). The Good Behavior Game in preschool classrooms: An evaluation of feedback. *Journal of Applied Behavior Analysis*, 52(1), 105–115. <https://doi.org/10.1002/jaba.500>
- Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in school psychology* (Vol. 4, pp. 251–288). Lawrence Erlbaum.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11(2), 203–214. <https://doi.org/10.1901/jaba.1978.11-203>
- Wright, R. A., & McCurdy, B. L. (2012). Class-wide positive behavior support and group contingencies: Examining a positive variation of the Good Behavior Game. *Journal of Positive Behavior Interventions*, 14(3), 173–180. <https://doi.org/10.1177/10983007114210>