All for one and one for all: The good inclusion game

Scott Graham¹ | Mickey Keenan¹ | Karola Dillenburger²

¹Department of Psychology, Ulster University, Coleraine, UK
²Centre for Behaviour Analysis, School of Social Science, Education and Social Work, Queen's University Belfast, Belfast, UK

Correspondence
Scott Graham.
Email: graham-s14@ulster.ac.uk

Funding information
Department for the Economy

Abstract

The Good Inclusion Game (GIG) is a class-wide education intervention employing interdependent group contingencies to increase inclusive behaviors between pupils with and without special educational needs (SEN). While previous research provided evidence of effectiveness when the class of children was split into two groups, it did not assess the effects of the GIG for one whole-class group or in the long term. The present study addresses both of these issues. Using an ABAB design, results show a one-team version of the GIG to be effective in increasing inclusive behaviors between pupils with and without SEN. Gains were maintained across time. Implications for inclusive education interventions and teacher training are discussed.

KEYWORDS
ABA, applied behavior analysis, good inclusion game, inclusive education, special education needs

1 | INTRODUCTION

Inclusive education has gained significant global momentum (Amor et al., 2018) since the Salamanca Statement (United Nations Educational, Scientific and Cultural Organisation & Ministry of Education and Science Spain, 1994). The expression "inclusive education" refers to the education of all children together, those with special educational needs (SEN) and their peers without SEN, in mainstream schools (Finkelstein et al., 2021; Weisel & Dror, 2006). In Northern Ireland, the term "SEN" is used to classify educational needs that are more substantial than those of the

The research was conducted in partial fulfilment of doctoral level study by the first author under the supervision of the second author.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). Behavioral Interventions published by John Wiley & Sons Ltd.
majority of pupils of a similar age (Education (Northern Ireland) Order 1996; SEN and Disability Order (Northern Ireland) 2005; SEN and Disability Act (Northern Ireland) 2016). Consequently, the SEN Code of Practice requires schools to adapt and provide the appropriate additional level of support to students who have SEN (Department of Education, 1998).

The notion of educational equality and equity is at the forefront of inclusive education, stating that all pupils have the right to the same educational content and experiences, with the appropriate supports in place where and when needed (Olson et al., 2016). Research suggests that inclusive education can result in positive social and academic outcomes for both pupils with and without SEN (Dalgaard et al., 2022; Ruijs & Peetsma, 2009). However, students with SEN still are more likely to be bullied (Humphrey & Hebron, 2015) and socially rejected by their peers without SEN (Broomhead, 2019; Garrote et al., 2017). Clearly, the inclusion and social participation of pupils with SEN in mainstream classrooms remains suboptimal (Garrote et al., 2017).

Given the potential academic and social challenges of maintaining an inclusive environment in schools, it is crucial for teachers to have access to evidence-based strategies that will enable them to cultivate a genuine inclusive atmosphere within the classroom and foster the skills and achievements of pupils with diverse needs. One relatively new and promising class-wide inclusive intervention is the Good Inclusion Game (GIG; Coyle & Dillenburger, 2019; Dillenburger & Coyle, 2019). The GIG is a variation of The Good Behavior Game (GBG), a class-wide intervention that has a large evidence base of effectiveness across populations (Donaldson et al., 2011; Keenan et al., 2000), behaviors (Hernan et al., 2019; Wright & McCurdy, 2012), and settings (Cheatham et al., 2017; Foley et al., 2019). Both the GIG and GBG typically consist of the following components: (a) the class is split into teams, (b) rules are presented, (c) feedback is delivered, (d) a criterion to earn a reward is established, and (e) rewards are provided contingent on meeting the criterion.

While in the GBG, the aim is typically to reduce disruptive behaviors such as talking out and leaving the seat (Barrish et al., 1969), in contrast, in the GIG, the focus is on increasing inclusive behavior (Coyle & Dillenburger, 2019; Dillenburger & Coyle, 2019). Accordingly, the GIG applies an interdependent group contingency in which the delivery of reinforcers is contingent on the children engaging in specified inclusive behaviors, namely “communicating” and “sharing.” These behaviors align with the Inclusion Index Indicator A1.1 “everyone is made to feel welcome” and Indicator A1.2 “students help each other” (Booth & Ainscow, 2011).

The developers of the GIG, Dillenburger and Coyle (2019), compared the GIG with the GBG for pupils with and without SEN in mainstream education settings and found that the GBG reduced disruptive behaviors but did not increase inclusive behaviors. In contrast, the GIG increased task-related inclusive behaviors, thus increasing the social participation of pupils with SEN. Moreover, the GIG also reduced disruptive behavior, despite the fact that these behaviors were not specifically targeted. However, although promising, more research is required. For example, the existing research only examined the effects of one brief 15 min GIG session compared to one brief 15 min GBG session in each of the nine participating classrooms. Thus, while showing external validity, the effects of the intervention on behavior over an extended number of sessions remain unknown. Furthermore, the question remains open, of whether or not the game has to be played in two competing teams. If the whole classroom can play as one team, this may increase the inclusivity of the intervention and streamline the application of the practice for teachers. Consequently, the objectives of the current investigation were to assess the effects of the GIG on inclusive interactions between pupils (with and without SEN) over several sessions and determine the game’s effectiveness when played as one whole-classroom team.

2 | METHOD

2.1 | Participants and setting

The present study was conducted in a mainstream school in a small town in Northern Ireland. The participating class consisted of 27 pupils (16 males and 11 females) aged between 6 and 7 years old. Four of the pupils were
identified as having SEN. Although all pupils provided assent and participated in the game, data were gathered only from pupils whose parents or guardians also had provided written consent. This meant that data were collected on a total of 17 pupils (nine males and eight females). The same male teacher taught the class throughout all sessions. Of the four pupils with SEN, one was designated as the target SEN pupil. This pupil was a 7-year-old male with a diagnosis of autism spectrum disorder and a mild learning disability. Despite not having an assigned classroom assistant, the teacher noted that he required more instructional support and displayed greater social difficulties (e.g., maintaining friendships) than the majority of the other pupils. For these reasons, the teacher selected this pupil as the target SEN pupil.

2.2 Target behavior

Two distinct categories of target behavior were identified: initiation and sustained interaction (Table 1). Initiation included (1) requesting information or an item, (2) offering or asking if others needed information or an item, and (3) positive or general comments. Data on these initiation behaviors were collected in aggregate. Sustained interaction behaviors consisted of verbal and motor responses that maintained the interaction. New initiations were counted if at least 5 s had elapsed between the last sustained interaction response and the next initiation response. For the overarching category inclusive behavior, data on initiation and sustained interaction were combined. Additionally, two teacher behaviors were measured: inclusive praise and point delivery. Operational definitions of teacher behaviors are detailed in Table 2.

### TABLE 1 Topographies and operational definitions of target behavior.

<table>
<thead>
<tr>
<th>Target behavior</th>
<th>Response topology</th>
<th>Operational definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation behavior</td>
<td>Requesting information or an item</td>
<td>Either the target pupil or their peer initiating an interaction with the other in relation to receiving an item or information related to the work task</td>
</tr>
<tr>
<td>Initiation behavior</td>
<td>Offering or asking if others need information or an item</td>
<td>Either the target pupil or a peer initiating an interaction with the other, relating to the other needing information or an item specific to the task</td>
</tr>
<tr>
<td>Initiation behavior</td>
<td>Positive or general comments</td>
<td>Either the target pupil or a peer initiating an interaction by making a statement related to the current task</td>
</tr>
<tr>
<td>Sustained interaction</td>
<td>Verbally responding appropriately to a peer</td>
<td>Either the target pupil or a peer emitting an appropriate verbal response to the other’s task-related verbal response and/or continued with verbal assistance</td>
</tr>
<tr>
<td>Sustained interaction</td>
<td>Physically responding appropriately to a peer</td>
<td>Either the target pupil or their peer physically following the instruction of the other or actively assisting the other</td>
</tr>
</tbody>
</table>

Note: Initiation behaviors and sustained interaction behaviors were later aggregated to produce a measure of “inclusive behavior.”

### TABLE 2 Operational definitions of teacher target behavior.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Operational definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusive praise</td>
<td>Any instance in which the teacher vocally acknowledges and delivers behaviour-specific praise to one or more pupils for engaging in one or more of the target inclusive behaviors</td>
</tr>
<tr>
<td>Point delivery</td>
<td>The teacher physically delivers a point</td>
</tr>
</tbody>
</table>
3 | MATERIALS

3.1 | The GIG script

The teacher paraphrases the following script:

Today, we are going to play The GIG—the goal is to help, share, and care for each other. The rules are:

1. You should show teamwork by talking to each other
   You are allowed to talk to help each other complete your classwork. This means you should work together to make sure each of you understands the work. For example, if you don't understand the work, ask one of your classmates, or if you see someone struggling, ask them if they want help.

2. You should share with others
   It is important that you share stationary with each other such as coloring pencils or things needed for the class task. This means you must be kind and share with others if they need or would like to borrow any of your materials. You are allowed to collect any pencils or things you or your classmates need to help with your work.
   The game will be played while you are doing your work. If I see you following the rules, I will place a ‘GIG point’ on the board. We will play the game as one big team. Therefore, you must work as a team to win. If you get five or more points on the board, the class will win a prize. The prize is (the teacher described the prize and when it could be accessed).

As recommended by Barrish et al. (1969), usually the teacher selects the prize to ensure the prize are available within the school setting (e.g., 5 min of extra recess; a night off from homework; candy). Given the nature of the selected prizes, this will result either in the immediate delivery of the reinforcer (e.g., candy) or a delay in the delivery of the reinforcer (e.g., foregoing homework or extra recess time later in the day).

3.2 | Observation recording sheet

A simple blank table was printed on an A4 sheet of paper. The table was divided into 20 s partial interval data recording boxes for each of the target behaviors. The 20 s observation interval was selected because it was found to be practical in the given environment while allowing for an adequate number of intervals. Frequency recording boxes were available in the table for point delivery (Cooper et al., 2020).

3.3 | Treatment integrity checklist

A treatment integrity checklist was adapted from Joslyn et al. (2020) and consisted of a list of teacher behaviors that were expected in the delivery of the GIG: (a) reading/paraphrasing the script, (b) showing the scoreboard, (c) putting up the GIG poster, (d) reminding the pupils of the rules, (e) announcing the start of the game, (f) starting the timer and placing it in his pocket, (g) delivering points, (h) giving behavior specific praise/awarding points, (i) playing the game for a maximum of 15 min or until the end of the class activity, (j) adding up the points and announcing if the team won, based on a pre-established criterion, (k) distributing rewards (if appropriate), and (l) taking down the scoreboard (the group avatar) and poster.
The researcher to hear from the peer interaction during the peer sessions, and the researcher collection, included the least for the least in the rules (Dillenburger et al., 2019). Dillenburger and her solitary, verbal, and pupils were also told for their target. They were asked to be appropriate for the pupils in the class. The teacher asked the pupils to engage in the positive in the classroom (ClassDojo, n.d.). The teacher also used “ClassDojo,” a free online tool for schools that has many functions, including assisting teachers with a positive behavior classroom management system (ClassDojo, n.d.). The ClassDojo was displayed on the classroom interactive whiteboard and each pupil was assigned their own avatar. The whiteboard was connected to an iPad®, which allowed the teacher to deliver points immediately for appropriate classroom behavior (e.g., attending to the teacher, providing a “good answer,” and helping peers). When a pupil obtained 100 points (performance criterion; Litow & Pumroy, 1975), they did not have to complete the next

### 3.4 Social validity questionnaire

A social validity questionnaire was developed including questions about intervention goals, procedures and outcomes (Wolf, 1978). All questions were presented in a multiple-choice format, featuring the options yes, no, and unsure (Table 3).

<table>
<thead>
<tr>
<th>Question</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you feel that “helping each other” is an important behavior to target?</td>
<td></td>
</tr>
<tr>
<td>Do you feel that “sharing” is an important behavior to target?</td>
<td></td>
</tr>
<tr>
<td>Do you feel the good inclusion game is useful?</td>
<td></td>
</tr>
<tr>
<td>Do you feel the good inclusion game is intrusive to class learning?</td>
<td></td>
</tr>
<tr>
<td>Do you feel the good inclusion game is easy to use?</td>
<td></td>
</tr>
<tr>
<td>Do you think the pupils enjoyed playing the good inclusion game?</td>
<td></td>
</tr>
<tr>
<td>Do you believe the good inclusion game helped prevent unwanted/disruptive behavior (e.g., off task speaking out and walking around)?</td>
<td></td>
</tr>
<tr>
<td>Do you believe the good inclusion game improved inclusion interactions between peers (e.g., communication and participation between peers on class tasks)?</td>
<td></td>
</tr>
<tr>
<td>Would you use the good inclusion game again?</td>
<td></td>
</tr>
</tbody>
</table>

### 4 PROCEDURE

#### 4.1 Baseline

Data collection was discreet as possible. The researcher (SG) was present in the classroom for at least 15 min prior to each session to facilitate familiarization for the pupils and reduce reactivity effects. During data collection, the researcher sat approximately 1 m from the target SEN pupil to ensure he could hear and record vocal interactions clearly.

Baseline observation sessions were conducted across a range of lessons that included active peer engagements. Lessons in which the pupils were expected to engage in solitary study were not included. Each observation session lasted a maximum of 15 min. Target interactions were recorded only if they involved the target SEN pupil (Coyle & Dillenburger, 2019; Dillenburger & Coyle, 2019).

During baseline sessions, the teacher was asked not to change any of his common classroom management methods. He used antecedent interventions, including verbal reminders telling children to follow classroom rules and specific phrases to direct pupil attention (e.g., “1, 2, 3, eyes on me”). He also used consequence-based interventions, such as occasionally delivering reprimands and vocal praise for disruptive and appropriate behavior, respectively. When the noise level in the class got too high, he stopped talking and waited for the pupils to quieten.

In conjunction with these strategies, the teacher also used “ClassDojo,” a free online tool for schools that has many functions, including assisting teachers with a positive behavior classroom management system (ClassDojo, n.d.). The ClassDojo was displayed on the classroom interactive whiteboard and each pupil was assigned their own avatar. The whiteboard was connected to an iPad®, which allowed the teacher to deliver points immediately for appropriate classroom behavior (e.g., attending to the teacher, providing a “good answer,” and helping peers). When a pupil obtained 100 points (performance criterion; Litow & Pumroy, 1975), they did not have to complete the next
homework assignment. During baseline sessions, the teacher gave points sporadically and sparsely, resulting in none of the pupils attaining the criterion performance.

4.2 Teacher training

The first author (SG) provided relevant training for the teacher, using written and verbal instruction, and modeling (Joslyn & Vollmer, 2020). Throughout the training, which lasted approximately 20 min, the teacher was encouraged to ask questions and comment. It was agreed that the ClassDojo® system would be used to record GIG points and that during the GIG, the logging of standard practice points would be paused (Lynne et al., 2017).

The teacher was trained to award GIG points according to a fixed-interval 30 s (FI30) with a 10 s limited hold (LH10) schedule of reinforcement. A timer was placed in the teacher’s pocket. The timer vibrated (no sound) for 10 s (LH10) after 30 s (FI30) had elapsed, during which time the teacher scanned the room for a target behavior. If he observed a target behavior during these 10 s, he delivered pupil and behavior-specific praise (e.g., “nice job asking for help Toby, that’s a point for the class”) and awarded a point. If no target behaviors were observed during the LH10, no points were awarded and the next FI30 began.

4.3 GIG

Lessons in which the pupils were expected to engage in solitary study are not appropriate or aligned with the objectives of the GIG. Thus, similar to baseline, intervention sessions were conducted across a range of lessons that allowed for pupil interaction. The GIG was played with the whole class constituting one group (as opposed to splitting the class into 2 competing groups as in the original GIG; Dillenburger & Coyle, 2019). After explaining the rules to the pupils, the teacher put a “group avatar” on the whiteboard, the GIG rules poster on the wall and started the game. GIG sessions lasted 15 min, except for Session 15, which finished after 11 min, when the pupils had completed the assigned task. At the end of each GIG session, the teacher announced whether or not the pre-determined criterion had been reached. Contingent on winning the game, the teacher praised the pupils and distributed the prize or advised when the prize would be delivered, if it was delayed (e.g., 5 min extra recess). If the criterion was not met, the teacher (using a neutral tone of voice) advised the class that they did not win the game and encouraged them to try again during the next GIG session. Finally, the teacher took down the GIG rules poster and removed the “group avatar” from the whiteboard. A total of 15 sessions were conducted, 2–3 days per week, with a maximum of 2 sessions daily.

4.4 Interobserver agreement

Interobserver Agreement (IOA) was calculated using data collected by the first author (SG) and one second observer. Two individuals served as second observers; one was a PhD researcher, and one was a Board Certified Behavior Analyst (BCBA). Both second observers were trained in the procedure using behavior skills training. The second observer sat at a location in the classroom where peer interactions involving the target SEN pupil could be seen and heard. IOA data were collected in 20 s intervals during at least 25% of all baseline and GIG sessions.

IOA was calculated by dividing the number of agreements between both observers by the number of agreements plus disagreements, multiplied by 100. IOA was 88.76% (range 81.8%–100%) for target SEN pupil behavior, 86.7% (range 75.8%–100%) for peer behavior, 80.4% (range 71.1%–100%) for combined pupil behavior, and 91.7% (range 75.8%–100%) for teacher behavior.
4.5  Treatment integrity

The treatment integrity checklist was completed for each GIG session. Treatment integrity was calculated by dividing the number of behaviors the teacher completed by the total number of expected behaviors on the checklist, multiplied by 100. Mean treatment integrity during GIG sessions was 95.8% (range 91.7%–100%).

4.6  Social validity

After the applied aspect of the study (playing the GIG) was completed, the teacher was asked to complete the short online social validity questionnaire. Similar to Joslyn et al. (2019), pupils evaluated the GIG in a class discussion format. They were asked if they liked the game, if they wanted to continue playing, and if there was anything they did not like about it.

4.7  Data analysis

An ABAB withdrawal design was used to evaluate the effectiveness of the GIG. After each session, data were graphed and analyzed visually for level, trend, and variability of data. The study was evaluated using the What Works Clearinghouse (WWC) specifications for evaluating single-case research designs (WWC, 2020). To meet the WWC design standards, the following design criteria need to be met: (a) the independent variable should be systematically manipulated, (b) IOA should be obtained in each phase and must be collected on at least 20% of all sessions within a phase, with a minimum threshold of at least 80% agreement, (c) a withdrawal design should have at least four phases, and (d) to meet the WWC standards without reservations, the first baseline phase must have at least six data points, and at least two phases per condition must have five or more data points per phase. However, an exception exists in which any phase with at least three data points and zero within-phase variability may also meet the standards without reservations. To meet the standards with reservations, there must be at least two phases per condition and at least three data points per phase.

As advised by the WWC standards, the effect magnitude of the intervention was measured using the non-overlap of all pairs (NAP) index. NAP is a reliable nonparametric method that compares each baseline data point with each subsequent intervention phase data point. The current study provides an overall weighted average NAP score, for total inclusive behaviors, across comparison phases. The developers of the NAP index recommend that NAP scores from 0.93 to 1 be interpreted as showing strong effects, scores from 0.66 to 0.92 as moderate effects and scores between 0 and 0.65 should be interpreted as weak effects (Parker & Vannest, 2009).

5  RESULTS

Figure 1 shows the percentage of intervals in which the target SEN pupil and his peers engaged in mutually inclusive interactions across baseline and GIG conditions. During the initial baseline period, inclusive behavior occurred at a consistently low level (range 2.2%–17.8% of the 20 s observation intervals; M = 11.9%). However, inclusive behavior increased with the introduction of the GIG (range 40%–62.2% of intervals; M = 51.1%). Withdrawal of the GIG led to greater variability in inclusive behavior compared to the initial baseline (range 8.8%–37.8% of intervals; M = 20.6%); however, an overall descending trend was observed. The reintroduction of the GIG produced another increase in inclusive behavior (range 46.7%–55.6% of intervals; M = 51.2%). The second GIG phase resulted in a similar level of inclusive behavior as the first GIG phase, however, a more stable trend was
observed. The overall weighted average NAP score across comparisons indicated that the intervention was highly effective at increasing inclusive interactions between the target SEN pupil and his peers (NAP = 1).

The top panel of Figure 2 shows the percentage of intervals in which the target SEN pupil engaged in initiation and sustained interaction behavior with his peers. During the initial baseline, both target behaviors occurred at low levels (initiation behavior range 2.2%–11.1% of intervals; M = 8.1%; sustained interaction range 0%–4.4% of intervals; M = 3%). The introduction of the GIG produced a slight increase in initiation behavior, although a descending trend was observed (range 6.7%–22.2% of intervals; M = 14.4%); a greater increase was observed in relation to sustained interaction behavior, with an ascending trend (range 17.8%–53.3% of intervals; M = 32.2%). A return to baseline conditions produced no significant change in initiation behavior (range 4.4%–13.3% of intervals; M = 7.8%), whereas an overall lower level and descending trend was observed in sustained interaction behavior (range 2.2%–31.1% of intervals; M = 13.3%). The reintroduction of the GIG was associated with a slight increase in initiation behavior (range 13.3%–24.2% of intervals; M = 18.2%) and an increase in the level of sustained interaction behavior (range 26.7%–35.6% of intervals; M = 30.1%). However, no data point was as high as the highest data point in the initial introduction of the GIG.

The bottom panel of Figure 2 depicts the percentage of intervals in which the peers engaged in initiation and sustained interaction behavior with the target SEN pupil. During the initial baseline, both behaviors occurred at low levels (initiation behavior range 0%–6.7% of intervals; M = 3.7%; sustained interaction range 0%–4.4% of intervals; M = 2.2%). When the GIG was introduced, there was a slight increase in initiation behavior, although a gradual descending trend was observed (range 13.3%–22.2% of intervals; M = 18.3%). In terms of sustained interaction behavior, the GIG produced a greater change in the level of responding, however, compared to initiation behavior more variability was observed (range 13.3%–44.4% of intervals; M = 29.4%). On return to baseline conditions, no significant change in initiation behavior was observed (range 2.2%–8.9% of intervals; M = 5%). Simultaneously, a reduction in the overall level and a decreasing trend in sustained interaction behavior was observed (range 0%–15.6% of intervals; M = 9.4%). When the GIG was reintroduced, there was a slight increase in initiation behavior (range 13.3%–26.7% of intervals; M = 16.1%). Sustained interaction behaviors returned to similar levels found in the first GIG phase, however, data were less variable (range 22.2%–33.3% of intervals; M = 28.3%).

Figure 3 depicts the teacher’s target behaviors. During the initial baseline phase, inclusive praise was low and stable (range 0%–2.2% of intervals; M = 0.7%). The introduction of the GIG was associated with an increase in the level of inclusive praise with fairly stable responding (range 17.7%–28.9% of intervals; M = 21.7%). The withdrawal of the GIG resulted in inclusive praise returning to baseline levels (range 0%–2.2% of intervals; M = 0.6%).
FIGURE 2 Initiation and sustained interaction behaviors for pupil with SEN and peer. BSL, baseline; GIG, good inclusion game; SEN, special educational needs.

FIGURE 3 Teacher target behavior. BSL, baseline; GIG, good inclusion game.
the GIG was reintroduced, inclusive praise returned to similar levels observed within the initial GIG phase (range 11.1%–26.7% of intervals; M = 18.6%).

As discussed above, points were delivered across all phases of the study, although on different schedules. During the initial baseline condition, the teacher delivered a low number of standard practice points across all sessions, (range 0–3 points; M = 1.3). When the GIG was introduced, an increase in points delivered was observed, with an initial spike followed by a declining trend (range 8–16 points across sessions; M = 11.5). Returning to baseline conditions produced a return to points delivered at similar levels as the initial baseline, with an outlier in Session 10 (range 0–8 points; M = 2.5). Reintroducing the GIG resulted in an overall increase in the number of points delivered (range 6–13 points; M = 8.3).

According to the results of the social validity measures, the intervention was well received by the teacher and pupils. The teacher responded positively to all questions. Out of the 17 pupils who participated, 15, including the target SEN pupil, expressed enjoyment of the game, while one pupil stated that they did not like the game and another was unsure (88.2%, 5.9%, and 5.9%, respectively). When asked if there was anything specific they did not like about the game, the pupil who had stated that they disliked it said it was “boring and too much work.” However, the other pupils expressed interest in continuing to play, with one saying “I love sharing; it makes people happy.” The pupil who was seated beside and therefore interacted most with the targeted SEN pupil stated “I love that we got to help each other.”

The current study met the VWC design standards with reservations (WWC, 2020). In other words, the study maintained a reasonably high quality in the design standards, but with some issues. Specifically, each phase consisted only of three to four data points. In the initial baseline phase, three data points were collected as the target SEN pupil missed school on one of the scheduled observations. This reflects the challenges in collecting data in a classroom setting and emphasizes the need to take into consideration the time constraints within these settings when evaluating study designs.

6 | DISCUSSION

The GIG is an effective intervention for increasing inclusive behaviors amongst pupils (with and without SEN) in mainstream classrooms (Coyle & Dillenburger, 2019; Dillenburger & Coyle, 2019). However, its effect had not been examined across time or using diverse group contingencies. The current study explored the effectiveness of the GIG across time and via a one-team group contingency approach, rather than the usual two-teams competing contingencies. Consequently, the present study contributes to the existing literature by showcasing how the GIG is a flexible, easy-to-use, and effective tool for teachers that can help maintain inclusive classroom behaviors.

The current study shows that the GIG proves to be a successful strategy when the whole class collaborates as a united team. Similar to research on the GBG by Donaldson et al. (2021), the current study suggests that the competitive contingencies generated between teams are not a necessary mechanism for the GIG to be deemed effective in increasing inclusive interactions. This is an important finding given the ethos underpinning the intervention. In other words, a one-team version of the GIG may contribute to the inclusivity of the intervention. Playing as one team means pupils do not have to be segregated into teams, thus potentially improving the inclusive atmosphere within and across the classroom. Indeed, in the process of conducting the current study, the first author (SG) met with the principal of the school to provide more information regarding the GIG and obtain informed consent. Whilst discussing previous research, specifically the team component, the principal was very supportive of the one-group approach and stated that she would have been uncomfortable with the idea of one group of pupils receiving a prize and the others not.

In addition to enhancing value-based cooperation and improving the inclusivity of the intervention, a one-team version of the GIG also offers practical advantages for the teacher. For example, the teacher does not have to spend time dividing the class into teams and ensuring points are delivered to the appropriate team. Another benefit of the
one-team version of the GIG is that all pupils receive the same consequence. In two-team versions of the game, if using activity reinforcers, the teacher would have to arrange different activities for the team that meets the criterion and the team that does not meet the criterion. In contrast, in the one-team version, the entire class gets access to the reinforcer if the criterion is met. Furthermore, if the game is played in teams, this may limit the types of reinforcers that can be utilized. For example, due to insufficient supervisory resources, the teacher may be unable to offer activity reinforcers, such as extra recess.

Besides increasing inclusive behavior among pupils, the current investigation also demonstrated that the GIG can have a positive effect on teacher behavior. Specifically, the GIG increased the amount of praise and points (i.e., reinforcers) the teacher delivered. Data regarding teacher delivery of “GIG points” deserves further elaboration. Visual inspection of the data in Figure 3 shows a downward trend in the delivery of “GIG points” across both GIG phases. This trend could suggest a correlation between point-delivery and opportunities to deliver points, thus inferring low levels of inclusive pupil behavior. However, data in Figures 1 and 2, and anecdotal class-wide peer-to-peer observations do not support this idea. Rather, due to having to attend to other tasks (e.g., classroom visitors and helping pupils with specific tasks and issues), the teacher seemed to have missed a number of opportunities to award points. Although clearly a limitation within the current study, the positive effects of the GIG on the pupil behavior despite relatively poor point scoring accuracy by the teacher (Joslyn & Vollmer, 2020), may demonstrate the flexibility and strength of the GIG in lively and unpredictable applied classroom settings.

The current data suggest that the GIG is not only an effective intervention but that it is also socially valid. Overall, the teacher and pupils agreed that the intervention was acceptable in terms of goals, application, and outcomes. In addition to the questionnaires and interviews, informal discussions between the first author (SG) and the teacher contributed to the social validity findings. During one of these discussions, the teacher stated that the school had always advised the pupils to help, share, and care for each other and to feel confident in asking their peers for help. However, whilst playing the GIG, the teacher noticed a distinct difference, stating: “They [the pupils] must feel more confident… or psychologically safe in asking for help now…as it’s a game, they must feel more secure in asking…” This statement highlights the importance of understanding and organizing reinforcement contingencies in order to target specific behaviors.

Clearly, the GIG not only encourages pupils to ask for and offer help but also generally increases positive social interactions. This is particularly important as the prevalence of SEN has increased dramatically, especially with regards to autism diagnosis (Centers for Disease Control and Prevention, 2024). Many of these pupils have social and communication challenges and mental health issues, typically limiting the benefits that an inclusive classroom setting can offer (Koegel et al., 2012). Thus, the GIG offers a valid addition to more elaborate social skill training programmes (Radley et al., 2016).

Although the current study demonstrated the effectiveness of the GIG as a class-wide, one-team intervention across time, it is not without its limitations. The current data suggests that increases in inclusive behaviors may be met with a “ceiling effect.” The study found that the highest percentage of inclusive behaviors occurred in Sessions 5 and 7, in which inclusive behaviors were observed in 62.2% of intervals. It is plausible to suggest that since pupils are expected to complete their own assignments while also assisting their peers, the highest data points in the current study may be close to the maximum and acceptable level of achievement (i.e., it would not be acceptable for a pupil only to assist a peer and neglect their own assignments). Such suggestions ought to be investigated in future studies.

Furthermore, although the principal of the school in the current study was apprehensive about the use of teams, previous research on the GBG suggests that teachers seem to prefer a two-team version (Donaldson et al., 2021). Similar findings may be true for the GIG, which would have practice implications. For instance, Tingstrom (1994) suggests teachers are more likely to implement interventions which they prefer and find acceptable. Therefore, more research is needed to examine the effects of teams within the GIG and the social competitive group contingencies involved.
Another limitation relates to the lack of generalization. Although desirable from an experimental perspective, the lack of inclusive behavior generalization during the second baseline phase is of social concern. When the GIG was not played and standard teaching practices were re-implemented, the increased levels of inclusive interactions that were present during the GIG were not maintained. Indeed, a lack of generalization and maintenance has been a concern in much of the GBG literature (Groves et al., 2022; Joslyn et al., 2019; Moore et al., 2022). However, it is important to remember that the current study was conducted over a relatively short period of time. With prolonged exposure to the GIG, social skills may generalize to the natural environment. Future research into the GIG should explore the variables responsible for generalization and maintenance of effects. If generalization does not emerge with prolonged exposure to the GIG, future studies may utilize methods to promote more explicitly generalization and maintenance effects (Dadakhodjaeva et al., 2020; Foley et al., 2019).

Another limitation of the current study is that the sessions were conducted at different times of the day and across a range of classroom lessons. However, this is not necessarily different from other studies. For example, Groves and Austin (2017) conducted daily sessions, that were separated by a minimum of two hours. This is similar to the current study. Additionally, Sewell (2020) conducted sessions across Math and Literacy lessons. Indeed, the same lessons were taught during the current study, within and across phases. Nevertheless, it is important for future research to attempt to control these possible extraneous variables.

A final possible limitation is the dense schedule in which the teacher was to provide feedback if inclusive interactions were observed. Indeed, such a dense schedule likely contributed towards instances of feedback omission. Regardless, it is important to note that the social validity measures indicated that the teacher found the intervention to be “useful” and “easy to use.” Nonetheless, future studies may wish to assess other feedback procedures.

It may also be beneficial for future research to investigate the effects of the GIG on disruptive behavior. Coyle and Dillenburger (2019) and Dillenburger and Coyle (2019) found that the GIG also functioned to reduce disruptive behavior relative to baseline levels, even though disruptive behaviors were not specifically targeted. Although anecdotal observations from the current study suggest that the same was true when the GIG was played in a one-team approach, in other words, disruptive behavior was reduced despite not being explicitly targeted, further research is necessary to validate these findings.

Additionally, future researchers may wish to empirically determine the effects of GIG across the classroom. A number of practical considerations played a role in the decision to collect data only on interactions involving the target SEN pupil in the current study. First, sessions were relatively short, and to make the data meaningful, there was a need to ensure an adequate number of intervals in which the target SEN pupil was observed. Secondly, many of the target behaviors consisted of vocal interactions related to the class assignment. Therefore, collecting experimental data on multiple interactions across the classroom would require the researcher and second observers to move around the room to ensure vocal interactions were assignment-related. This could potentially disturb the teacher doing his job and pupils engaging in their tasks. For these reasons, the decision was taken to focus solely on interactions involving the target SEN pupil, allowing for more reliable and meaningful data to be collected. However, alternatively, future researchers could systematically record peer-to-peer interactions occurring in close proximity to their seated observation position. This could be achieved using a fixed-order data collection system alternating between the target SEN pupil and specific peers, similar to Bohan and Smyth (2022) and, Groves and Austin (2020). While this would not measure class-wide interactions, it would provide a sample of peer-to-peer interactions. It is important to note, though, that this would require longer sessions to ensure an adequate number of SEN-peer and peer-to-peer interaction intervals are recorded.

Finally, this study proposes that future research explores the effects of peer support. The contingencies of the GIG promote peer-tutoring and thus can transform the classroom into a peer-tutoring environment (Koegel et al., 2012). Previous research also found that the GBG can improve academic performance (Weis et al., 2015). Similar outcomes may be found with the GIG, but this remains to be explored. This would have great social significance as it would confirm that the GIG has the potential to address many concerns that mainstream teachers...
have regarding teaching in inclusive settings (Arrah & Swain, 2014; Braunsteiner & Mariano-Lapidus, 2021; Øen & Johan Krumsvik, 2022).

The question remains, why it is important to give this intervention the new name, the ‘GIG?’ Could it not simply be referred to as another version of the GBG? Of course, the same question could be asked of the “Caught Being Good Game” (CBGG; Bohan & Smyth, 2022), the “Timely Transition Game” (Yarbrough et al., 2004), or the “PAX GBG®” (Praxis Institute, 2021), and others, as all these games are iterations of the GBG that involve interdependent group contingencies to target specific classroom behaviors. Yet, different names are important as they allow teachers and researchers to differentiate efficiently between the objectives of the intervention. For example, the GBG is typically used to reduce disruptive behavior within the classroom and utilizes both, conditioned punishment and reinforcement contingencies (Brennan et al., 2024). The CBGG is used to increase on-task classroom behavior using conditioned reinforcement contingencies (Brennan et al., 2024). The GIG focuses on improving interactions between pupils with SEN and their peers within inclusive education classrooms using conditioned reinforcement group contingencies. While all of these games are free to be used by teachers anywhere in the world, the PAX GBG® is a commercialized version of the GBG.

In conclusion, the empirical data reported here support the effectiveness of the GIG in increasing inclusive interactions between pupils with and without identified SEN over a prolonged number of sessions. Additionally, the present study contributes to the literature regarding the effectiveness of the GIG when played in one class-wide team, improving inclusivity and streamlining the application. Since the Salamanca Agreement, there has been a notable increase of pupils with SEN attending mainstream classrooms. Still, these pupils experience more bullying, are often ostracized and excluded, and usually require more support with class assignments (Dillenburger, 2012). These challenges have not been met with adequate teacher training (Dillenburger et al., 2016). Utilizing peer-to-peer pupil support, the GIG, especially in a one-team format, is an easy-to-use, effective tool for teachers faced with the challenges of creating inclusive classrooms. Furthermore, given its effectiveness in the mainstream classroom, it holds promise for other settings, such as sports, arts, music, and youth clubs more generally.

ACKNOWLEDGMENTS
This research was funded by a Strategic Department for the Economy, Northern Ireland (DfE) PhD Scholarship. The authors would like to acknowledge Dr Dean Reid and Rachel Lewendon for their assistance in collecting data for this project.

CONFLICT OF INTEREST STATEMENT
We have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT
The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT
Ethical permission for the study was granted by the Ulster University Research Ethics Committee.

INFORMED CONSENT
Informed consent was gathered from all participants.

ORCID
Scott Graham https://orcid.org/0000-0002-9523-2431
Mickey Keenan https://orcid.org/0000-0002-5579-9169
Karola Dillenburger https://orcid.org/0000-0002-3410-5949
REFERENCES


---

**How to cite this article:** Graham, S., Keenan, M., & Dillenburger, K. (2024). All for one and one for all: The good inclusion game. *Behavioral Interventions, e2048*. https://doi.org/10.1002/bin.2048