Investigating Game-based Interventions for Classroom Management in Mainstream School Settings

Clare Bohan, BSc. (Hons) in Psychology

Thesis submitted in partial fulfilment for the award of

Doctor of Philosophy (PhD)

School of Psychology

Dublin City University

Supervisors:

Dr. Sinéad Smyth, School of Psychology, Dublin City University

Dr. Claire McDowell, School of Psychology, Ulster University

April 2021

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy (PhD) is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: ID No.: 16213015 Date: 23/04/21

Acknowledgements

This PhD journey could not have reached completion without the support of several people whom I wish to express my gratitude to. First, my primary supervisor, Dr. Sinéad Smyth, who gave me the initial chance to commence this endeavour and has supported me every step of the way. There were many roadblocks over the course of the four years, however Sinéad was a constant source of inspiration and encouragement for which I am infinitely grateful. I also wish to extend my thanks to my secondary supervisor, Dr. Claire McDowell for her insight on pieces of writing and support. Thank you to my independent panel member, Dr. Veronica Lambert for ensuring timely progression during my annual reviews. We are incredibly fortunate in the School of Psychology in DCU to have such an exceptional team of supportive staff. I would like to thank Dr. Lorraine Boran for providing critique and direction during my transfer examination and Patrick Boylan for technical support in securing relevant programmes needed for completion of a thesis. To the wider staff across the school, for the smile on the corridor, the quick chat in the kitchen and for providing a constant feeling that 'you can do this'; thank you. Having spent 8 years studying in DCU, I've watched the School of Psychology form and grow and it will always hold a special place in my heart.

Completing such a large piece of research could not have happened without the voluntary efforts of some wonderful academics who assisted on many elements. Thank you to Dr. Ronda Barron, a fantastic second reviewer and source of support and assurance. Thanks also to Shauna Glennon and Dorotea Dejanovic who assisted in other tasks as part of the review. A special thanks must also be extended to the brilliant team of observers who volunteered to help out with data collection in schools; you shortened the road for me on many a walk or drive to schools and helped this piece of research be as integral as possible.

I must extend a huge thank you to the principals, teachers and students who took part in this research- without your interest in the process and dedication to seeing it through, this thesis would not have been possible. I have great admiration for the work you do.

There are too many post graduate students to name in DCU, particularly the H101E gang, who have become such treasured friends over the past 4 years. Thank you all for the tea, the laughs, the walks to the shop (for ice cream), the laps of the park, the virtual catch ups since we left the office and the constant sense of community and support. I can't wait to see the wonderful things the future holds for you all!

To my mother and father, Mary and Bernie- there was never anything I felt I could not do, and this achievement is a direct consequence of your love and support. Thank you both for all you have done for me- for listening when I needed a listening ear, for being a trusted source of life advice and having my back, no matter what educational path I chose to follow. To Lisa, Lee and the boys- thank you all for the support over the last two years in particular. You are such wonderful selfless people and I am so grateful for the time you allowed me to be a lodger. Thank you to Rhona, Niamh and Brendan- for providing me with endless snippets of advice, laughs when I needed them and for always kept me grounded. To my godmother Helena, for always making me feel like one of the most special people in the world and having faith that I would get to the finish line, thank you. Finally, thank you to Seán. There are not enough words to convey my most heartfelt appreciation for all that you have done for me as I navigated this journey. I'm sure at times it felt like you had signed up to do a PhD too! I will be forever grateful for your patience and unwavering love and support.

Table of Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

Abbreviation	Explanation
ABA	Applied Behaviour Analysis
AEB	Academically engaged behaviour
BIRS	Behavioural Intervention Rating Scale (Elliot & Von Brock Treuting, 1991)
CBGG	Caught Being Good Game
CBGG-d	Caught Being Good Game with Delayed Feedback
CBGG-i	Caught Being Good Game with Immediate Feedback
CIRP	Children's Intervention Rating Profile (Witt & Elliott, 1985)
DB	Disruptive behaviour
DEIS	Delivering Equality of Opportunity in Schools: A status grated to some Irish schools which aims to bridge gaps in disadvantage between schools.
GBG	Good Behaviour Game
IOA	Interobserver Agreement
IRP-15	Intervention Rating Profile-15 (Martens et al., 1985)
JABA	Journal of Applied Behavior Analysis
MD	Motor disruption
OOS	Out-of-seat behaviour
PLS	Plain Language Statement
SNA	Special Needs Assistant
VD	Verbal disruption
WWC	What Works Clearinghouse

Abstract

Thesis Title: Investigating Game-based Interventions for Classroom Management in Mainstream School Settings

Student: Clare Bohan

Game-based behavioural interventions provide a potential solution to classroom management issues faced by teachers at all levels. Schools typically focus on the encouragement of positive behaviour in modern behaviour management and positive game-based interventions hold promise in addressing behaviour issues in the classroom. The current thesis was concerned with collating research on game-based interventions and evaluating an established positive intervention, the Caught Being Good Game (CBGG) in Irish mainstream classrooms.

Following an introductory chapter, Chapter 2 detailed a systematic review that collated research on game-based classroom management interventions in mainstream education. Findings indicated that many game-based classroom management interventions exist, with most of the research being conducted in the United States of America. Effect size calculations and a series of meta-analyses suggested that game-based interventions had moderate to large effects on classroom behaviour. This review led to the identification of the CBGG as an intervention warranting evaluation in Irish schools.

The CBGG was evaluated across Chapters 3-5, focusing on procedural variations and diverse contexts. In Chapter 3, the CBGG, both with and without visual feedback, was found to be effective in targeting academically engaged behaviour (AEB) and disruptive behaviour (DB) across three classes of young adolescents. In Chapter 4, the CBGG effectively targeted AEB and DB across an all-male middle primary school class, and two individual target students within the class. Chapter 5 outlined the effectiveness of the CBGG in targeting AEB and DB in a mixed-sex, early primary school class and with two individual students in that class. The CBGG remained effective as the schedule of reinforcement was thinned - a potentially useful method for reducing workload for the teacher. In Chapter 6, consideration was given to the findings in the context of the existing literature, strengths and limitations of the programme, and implications for future research and practice.

Chapter 1: Introduction

The research reported in the current doctoral thesis comprises a systematic review of the literature on game-based classroom interventions as well as a series of studies that sought to investigate one type of game-based classroom management intervention - the Caught Being Good Game (CBGG), in various different school settings. The purpose of the systematic review in Chapter 2 was to address the need for a review of the literature around the application of game-based interventions for classroom management in mainstream classroom settings. In particular, it aimed to identify the game-based interventions which have been applied, how they have been applied, the types of behaviour they have targeted and assert whether they have been effective in producing improvements in classroom behaviour. The empirical research studies in Chapters 3-5 evaluated the impact of the CBGG with Irish students from senior infants (aged approximately 5-6 years) to first year (aged approximately 12-13 years), maintaining a focus on the impact of procedural variations during the game.

The Irish School System

The school system is an integral part of Irish society, catering for 939,166 students and 66,961 full-time teaching staff in 2019/20 alone (Department of Education and Skills, 2020). For 183 days a year at primary level and 167 days a year at second level, Irish students and teachers attend school. The school day lasts 5 hours and 40 minutes for primary school students (with some junior classes spending slightly shorter hours in school) and at least 6 hours for second level students, meaning Irish students are spending in excess of 1000 hours in school each academic year. The hours students spend in school incorporate breaks, assemblies, excursions and extra-curriculars, but most of the time is spent in the classroom with their peers and class teachers. Given this extensive amount of time dedicated to education, schools are under pressure to contribute solutions to youth-based societal issues, producing well socialised, motivated young people who are ready to flourish as adults (Walker, 2004). Much of this pressure falls on teachers, who work face-to-face with students within the realm of their classrooms on a day to day basis. One of the key challenges faced by teachers in Ireland and worldwide, is disruptive and challenging student behaviour, which can contribute to teacher job dissatisfaction (ASTI, 2018) and burnout (Aloe et al., 2014). Although it is teachers who must manage and deal with student disruption, students themselves face potential negative outcomes, with misbehaviour being linked to future delinquency and criminal conduct (Smith, 2006), and poorer academic outcomes (Malecki & Elliott, 2002). It is agreed by both students and teachers that too much time is spent on the management of disruption in the classroom and that it is minor incidences of disruption happening at high rates which is the cause of much time-wasting (Infantino & Little, 2005). Time-wasting subsequently leads to loss of instructional time, putting well-behaved students at a disadvantage and placing teachers under time pressure to cover the curriculum. It is therefore evident that classroom management and more specifically, student behaviour management, is at the core of successful educational processes.

Classroom Management

While varying definitions of classroom management exist, the term is generally used to refer to actions taken by teachers with the intention of establishing order in their classroom, fostering student engagement, or enhancing student co-operation (Emmer & Stough, 2001). It is not a unitary concept, but rather comprises a range of elements and interrelated teacher decisions from the seating plan, to the classroom rules and daily routines. For example, student co-operation and engagement could be successfully fostered by considering the physical arrangement of the classroom carefully, in order to minimise crowding and distraction (Simonsen et al., 2008), but could also be influenced by implementing a behaviour management strategy such as a group contingency (see Maggin et al., 2012; 2017 for reviews). A combination of interconnecting decisions made by the teacher fosters a positive learning environment. Historically, discussions on classroom management date back to 1907, with a textbook published by Bagley (1907) serving as a primary resource on the topic (Brophy, 2006). Brophy (2006) outlines the history of classroom management since that 1907 publication, noting that there was little work in the early part of the 20th century, and empirical studies did not begin to gather pace until the middle decades. Incorporated in this uptake in research on classroom management, was a focus on leadership styles, group climate and punitive versus non-punitive approaches (e.g., Kounin & Gump, 1961). Behavioural approaches also became popular around this time, with the second half of the 20th century seeing a surge in behavioural classroom research (Landrum & Kauffman, 2006).

Behavioural Approaches in Classroom Management

Behavioural approaches to classroom management are derived from the behaviourist perspective in psychology, a perspective with a focus on overt, observable behaviour (termed ‘responses’) and environmental influences on this behaviour (termed ‘stimuli’). John Watson was an early pioneer of this approach who stated that a theoretical goal of psychology as a behaviourist views it is the ‘prediction and control of behaviour’ (Watson, 1913, p. 158), and put forward an early form of behaviourism, the stimulus-response (S-R) approach. B.F. Skinner, another pioneer of the behaviourist approach, proceeded to name a new science, the experimental analysis of behaviour, where he introduced ‘operant behaviour’. Operant behaviour is influenced by the stimulus changes which have occurred after the behaviour in the past, rather than being brought about by stimuli preceding the behaviour or antecedent stimuli (Cooper et al., 2007). He emphasised that that operant behaviour ‘operates upon the environment to generate consequences’ (Skinner, 1953, p. 65). Based on years of laboratory research, Skinner introduced operant conditioning, that is, the strengthening of an ‘operant’ (i.e., a class of responses), making a response more probable or frequent in the future (Skinner, 1953). Operant conditioning also encompasses incidences where responses become less probable or frequent in future (Cooper et al., 2007). There are two key behavioural principles at the core of operant conditioning: reinforcement and punishment. Reinforcement refers to the increased future frequency of a specific response when it is followed closely by a) the presentation of a stimulus termed a ‘positive reinforcer’ (i.e., positive reinforcement) or b) the removal or termination of a stimulus, termed a ‘negative reinforcer’ (i.e., negative reinforcement; Cooper et al., 2007; Skinner, 1953). Conversely, punishment refers to the reduced future frequency of a specific response following the presentation (positive punishment) or removal/termination of a stimulus (negative punishment; Cooper et al., 2007). Operant conditioning can therefore be understood as a three-component procedure, with an antecedent stimulus (referred to as a discriminative stimulus or S^d), a response (or behaviour) and a consequence (reinforcement or punishment). Although much of the early work of Skinner was conducted with animals in laboratories, in the middles decades of the 1900s, the field of applied behaviour analysis (ABA) began to emerge, applying the principles of behaviour to humans (Baer et al., 1968; Skinner, 1953). It was not long before the principles were applied in classrooms with students (early examples will be discussed later). To briefly demonstrate the application of these principles, Table 1.1. outlines how they may be applied in a classroom setting at their most basic level to target various student responses.

Table 1.1.

Behavioural Principles: Examples in the Classroom

Behavioural Principle	Example in the Classroom	Behaviour/Response Targeted[a]	Stimulus
Positive Reinforcement	When the teacher observes Alex raise his hand appropriately, she provides verbal praise and a sticker for his journal	Hand-raising	Verbal praise + Sticker
Negative Reinforcement	When the teacher observes Alex complete all of his worksheet, she gives him a night off homework	Worksheet completion	Homework
Positive Punishment	When the teacher observes Alex hit his peer, she gives him extra homework for that night	Hitting	Additional Homework
Negative Punishment	When the teacher observes Alex hit his peer, she knocks 5 min off his outdoor time at lunch	Hitting	Outdoor time

[a]Note that although these behaviours are targeted in these examples, reinforcement/punishment has not occurred unless the future frequency of that behaviour increases/decreases.

The importance of maintaining an understanding of operant conditioning and specifically the principles of reinforcement and punishment cannot be understated. They form core features of behavioural approaches to classroom management and in applying them systematically, they can have positive influences on student behaviour. Positive strategies, with a focus on positive reinforcement where possible, have been central to discussions around classroom management. Punishment, while effective in certain situations, can be a controversial strategy. Some of this controversy may stem directly from the negative connotations associated with the word ‘punishment’ as it is used in lay-man’s terms. For example, corporal punishment is still prevalent worldwide, and is associated with several problems for children, such as mental and physical health and behavioural and academic problems (Heekes et al., 2020) and it has banned in Irish schools since 1982 (Department of Education, 1982). Punishment may be used effectively and safely if ethical considerations are adhered to, for example, ensuring the procedure causes no physical or emotional harm, that it is the least restrictive procedure available and that it aligns with the individual’s immediate needs (Cooper et al., 2007). In classroom settings however, for low-level disruptive behaviours, a less restrictive alternative to punishment procedures should first be trialled before punishment is considered (Cooper et al., 2007). The potential for use of a positive classroom management intervention, focused on encouraging positive behaviour (the CBGG), in place of a similar intervention which draws teacher attention to disruptive behaviour (the Good Behaviour Game), will be discussed in further detail later in this chapter. Many of the discussions of behavioural applications in classroom management from here forward refer to applications in classrooms in the United States of America (USA). It is therefore pertinent at this point to outline that within the USA’s school system, Kindergarten is approximately equivalent to an Irish senior infants class (approximate age 5-6 years), first grade (approximate age 6-7 years) is approximately equivalent to first class, second grade is approximately equivalent to second class and so on.

Evaluating Behavioural Approaches

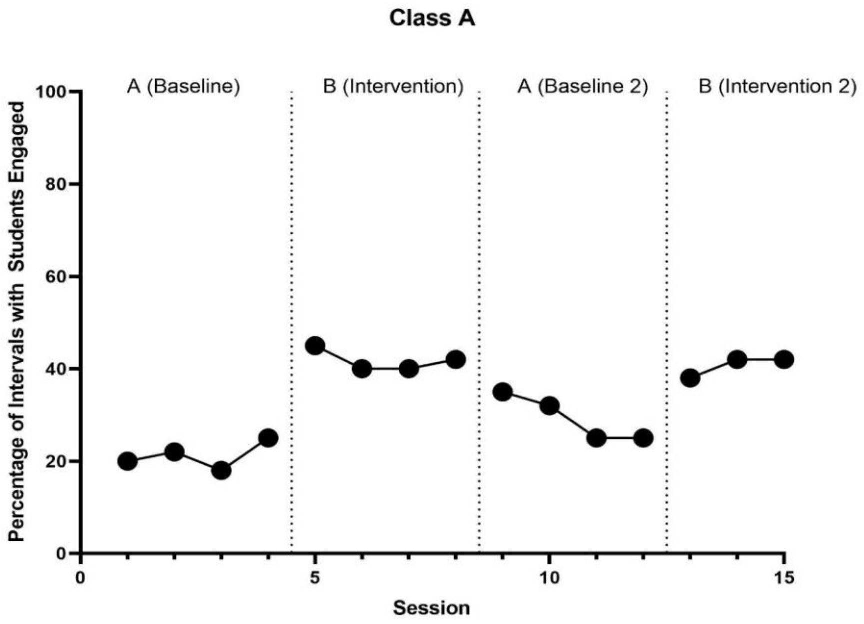
The rise in behavioural classroom management approaches in the middle of the 20th century coincided with the emergence of Applied Behaviour Analysis (ABA) which is a science maintaining a focus on observable behaviour and demonstrating functional relationships between events and the occurrence or non-occurrence of behaviour (Baer et al., 1968). In evaluating behavioural approaches to classroom management, single-case research designs are common. These are design types which involve the continuous monitoring of a target behaviour over time generally beginning with a baseline phase which serves two important functions: 1) to describe the behaviour under study to assert the existing level of performance and 2) to allow the researcher or practitioner to predict future levels of behaviour should intervention not occur (Kazdin, 2011). Single-case research designs have many distinct benefits in applied settings, which Kazdin (2011) outlines in detail, namely: 1) they provide useful scope for evaluation in applied settings where large scale between-group studies are not possible, 2) feedback is ongoing during intervention implementation, allowing the researcher or practitioner to continually assess behaviour rather than waiting for the results of a post-test, 3) the ability to quickly assess the generality of a treatment/intervention to other groups or settings and, 4) the single-case research design focuses on the individual or small group meaning the results are of direct and applied importance to them. It is important that the researcher undertaking to use these methods, has a clear understanding of their appropriateness for use in order to ensure maintenance of scientific rigour and integrity. For example, one important component when implementing these design types, is the collection of interobserver agreement (IOA) data during behavioural observations. Two data collectors observe the same behaviour at the same time, and to assert that the data collected are objective and reliable, the two observers should reach a certain level of agreement, that is, record the same data (Mitchell, 1979). Single-case research designs have received attention in recent years, with the Institute of Education's What Works Clearinghouse (WWC) division outlining design standards which should be met in their implementation (WWC, 2017; 2020). These design standards will be discussed in more detail in Chapter 2, but their very existence demonstrates the recognition and respect single-case designs have harboured in recent years.

Throughout the current thesis, single-case research designs will be reviewed in detail during a systematic literature review (Chapter 2) and will also be implemented across four empirical studies (Chapter 3-5). Two of the most common types of single-case research design are the reversal/withdrawal design (ABAB designs) and the multiple baseline design. Before discussing empirical examples of ABA in classroom management, it is important to briefly describe these design types.

Reversal/withdrawal Design (ABAB Design). Kazdin (2011) described the ABAB design as a design type where a baseline phase (i.e., no intervention phase, usually termed the 'A' phase) is alternated with phases where an intervention targeting the observed behaviour is in place (i.e., the 'B' phase). This pattern of baseline-intervention can be repeated and, to demonstrate experimental control, this must be done at least twice (i.e., ABAB). An AB iteration is not enough to rule out threats to external validity (Kazdin, 2011). There are many variations on this design type. For example, one may wish to evaluate the effects of more than one intervention type and therefore implement an ABAC, ABCABC or ABACABAC sequence of phases. The 'C' phases represents another intervention type. An example of an ABAB design applied to improve academically engaged behaviour of a class of students (Class A) is illustrated in Figure 1.1. (note, Figure 1.1. is composed of fictitious data for illustrational purposes).

Figure 1.1.

An Example of a Reversal/Withdrawal Graph (ABAB) Demonstrating the Effects of an Intervention on Student Engagement in Class A

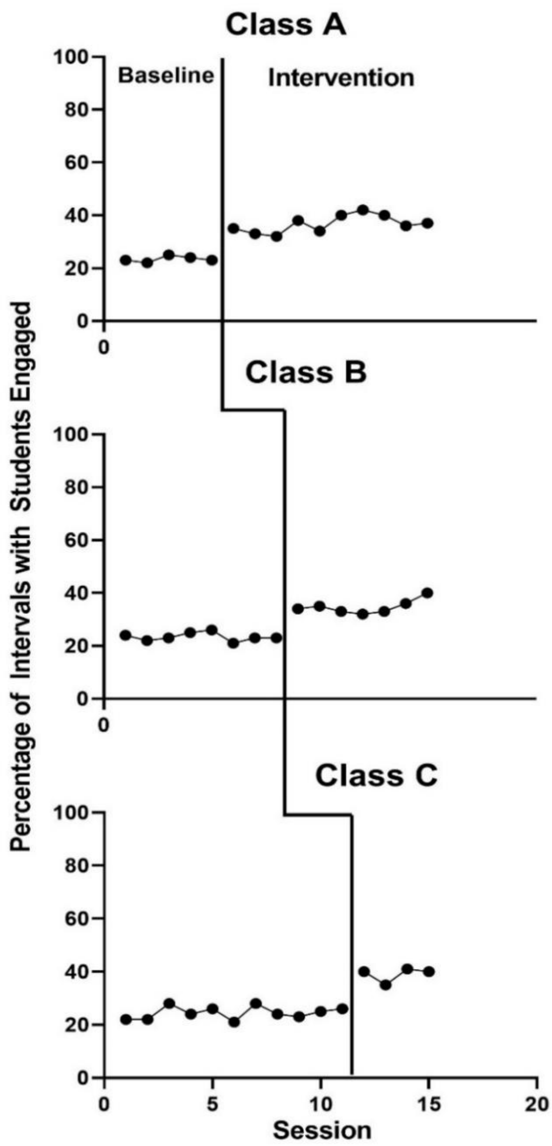


A very recent example of a reversal/withdrawal design in the classroom management literature was in a study by Ford et al. (2020), who evaluated a version of a popular intervention, the Good Behaviour Game (GBG; a classroom management intervention which will discussed in further detail later in this Chapter) across three classrooms. In their study, baseline levels of disruptive behaviour and academic engagement were monitored across five observation sessions in each classroom, which allowed the researchers to assert that improvements were warranted (i.e., disruptive behaviour was occurring at a high rate and engagement at a low rate). The GBG was subsequently implemented for five to seven observations sessions in each classroom, and improvements were evident in that disruptive behaviour systematically decreased and academic engagement systematically increased. Ford et al. then put a withdrawal phase in place and behaviour reverted to near baseline levels, before the GBG was implemented in a final phase and behaviour improved again. The researchers therefore systematically demonstrated that the GBG was the reason for the positive changes in behaviour.

Multiple Baseline Design. Multiple baseline designs involve examination of behaviour across several baselines (generally three). These baselines refer to different behaviours, settings, or participants. For example, if conducting a multiple baseline across behaviours design across three distinct behaviours, one would collect baseline data on those three behaviours. When the baselines demonstrate stability across the three behaviours, the intervention is implemented in targeting one behaviour, while baseline data continues to be gathered for behaviours two and three. When the intervention demonstrates effectiveness with behaviour one, the intervention is implemented with behaviour two, and subsequently behaviour three (Kazdin, 2011). This staggered implementation allows for demonstration of experimental control. An example of a multiple baseline design applied to improve academically engaged behaviour across three classes (Class A, B and C) is illustrated in Figure 1.2. (as with Figure 1.1., note that Figure 1.2. is composed of fictitious data for illustrational purposes).

Figure 1.2.

An Example of a Multiple Baseline Across Participants Groups Graph Demonstrating the Effects of an Intervention on Student Engagement Across Classes A, B and C



Using a multiple baseline across settings design, Floress et al. (2017) implemented the Caterpillar Game, a classroom management system aimed at reducing disruptive behaviour, with a first-grade class (ages ranged from 6-8 years). The intervention was put in place across three distinct subject settings: Carpet time, Seat Work Activity 1 and Seat Work Activity 2. Baseline levels of disruptive behaviour were monitored across these three subject settings initially before the Caterpillar Game was then put in place during Carpet time, while no intervention was in place during the Seat Work activities. When an improvement in disruptive behaviour was evident during Carpet time, the Caterpillar Game was put in place during Seat Work Activity 1, while baseline data collection continued in Seat Work Activity 2. Finally, when improvements were evident during Seat Work Activity 1, the intervention was put in place during Seat Work Activity 2, such that, by the end of the study, the Caterpillar Game was in place in all three settings. The researchers therefore systematically demonstrated that the improvements in disruptive behaviour were due to the Caterpillar Game rather than some other environmental variable.

Other Single-case Research Designs. Although reversal/withdrawal and multiple baseline designs are the most common in the behavioural literature, other design types exist which have been implemented in classroom management evaluations. The changing criterion design usually begins similarly to a reversal/withdrawal design, with a baseline phase. This phase is followed by a phase where the target individual or group must exceed or remain below a specific criterion of performance. Once the criterion has been met across a number of sessions, the criterion is increased/decreased, depending on the desired direction of the target behaviour. This design type is useful when targeting behaviours for gradual increase or reduction, for example, caffeine intake (e.g., Foxx & Rubinoff, 1979). It has been used to demonstrate the efficacy of classroom management interventions which involve students working towards a criterion, such as a certain amount of points or fouls (e.g., Beeks & Graves, 2016). Another design type, the alternating treatments design, is suitable where two or more distinct interventions are to be compared with each other. A baseline phase is usually (but not always) put in place to assert the state of behaviour before the interventions are put in place. This is followed by a phase where rapid alternation of the treatments under study takes place (Cooper et al., 2007). It is a useful design type in quickly identifying the more effective of two treatments and does not require a withdrawal phase. Hernan et al. (2019) recently used this type of design to compare the efficacy of a basic intervention to discourage phone use during class (Clear Box) with a more intensive intervention (Clear Box + GBG), finding that the more intensive intervention was much more effective in producing increases in engagement, decreases in off-task behaviour and decreases in phone usage during class.

Early Evaluations of Behavioural Interventions in Classroom Management

As identified earlier, the foundation of ABA as an applied science of behaviour targeting socially significant behaviour for improvement (Baer et al., 1968), coincided with a proliferation of behavioural studies relating to classroom management. The Journal of Applied Behavior Analysis (JABA) was established in 1968, as an outlet for disseminating ABA research. In the very first volume and issue of the journal (published on the 1st of March 1968), three of the eight original studies published were related to classroom management issues (Bushell et al., 1968; Hall et al., 1968; Thomas et al., 1968). Hall et al., (1968) observed increases in the study behaviour of six students (one first grade and five third grade students) during phases where the teacher provided brief social reinforcement (i.e., attention, a brief physical gesture such as a pat on the shoulder) contingent on study behaviour. This was demonstrated using an ABAB reversal design, where the B phases referred to phases where the teacher provided social reinforcement contingent on study behaviour. Thomas et al., (1968) manipulated the approval statements of a middle primary class teacher by having her withdraw approval statements altogether and only use contingent disapproval statements to control the classes' behaviour. Disruptive behaviour occurred at a higher rate during these 'no approval' phases when compared to baseline phases where the teacher could use approval statements.

In another paper in that issue of JABA, Bushell et al., (1968) outlined the application of a group contingency intervention with a preschool class. Students could earn tokens while working appropriately during independent study periods, Spanish instruction and study teams sessions (sessions where students worked in pairs and taught each other skills). These tokens could be traded for a special activity. The authors observed higher rates of study behaviour among the preschool class during sessions where the special activity was contingent on appropriate study behaviour. This application is an example of an independent group contingency, a group contingency whereby all members of the group are subject to the same response contingencies and only those meeting the criteria independently are eligible for the specified reward (Litow & Pumroy, 1975). There are two other types of group contingency: dependent and interdependent. All three types of group contingency will be discussed in further detail later in this chapter.

The three 1968 studies discussed provide an insight into the effectiveness of positive reinforcement in the classroom and also highlight the focus on classroom management in the ABA literature from the inception of the flagship journal. It is also worth highlighting that the Bushell et al., (1968) paper provided an early example of two of the most common behavioural strategies for classroom management: the token economy and the group contingency. These two strategies have evidently been prevalent since the middle of the previous century and are still subject to empirical investigation today. Further consideration will now be given to both approaches.

Token Economy Interventions in the Classroom Management Literature

The token economy was briefly considered earlier in this chapter when outlining the research of Bushell et al. (1968). A token economy is an approach whereby tangible generalised conditioned reinforcers (i.e., tokens) are provided contingent on a target behaviour, which can later be traded for backup reinforcers (Kazdin & Bootzin, 1972). Generalised conditioned reinforcers are reinforcers which are associated with a wide range of primary reinforcing stimuli, such as money which can be exchanged for a range of primary reinforcers (Skinner, 1953). Advantages in using this type of reinforcer include: 1) bridging the delay between a target behaviour and reinforcement, 2) reinforcement can occur at any time, 3) the reinforcers are less subject to satiation and 4) the tokens can provide similar reinforcement for those who have different preferences when it comes to primary reinforcers (Kazdin & Bootzin, 1972), all of which make them potentially desirable for use in a classroom setting. One of the earliest examples of token economy application in a mainstream classroom context, was in a study by O’Leary et al., (1969). Seven students were observed in a second-grade class across eight phases: 1) Baseline, 2) Classroom Rules (rules placed on the blackboard), 3) Educational Structure (structure of the period enhanced), 4) Praise Appropriate Behaviour and Ignore Disruptive Behaviour, 5) Tokens and Back-up Reinforcement (points administered four times throughout the period which could be exchanged for back-up reinforcers), 6) Withdrawal, 7) Token Phase Reinstated and 8) Follow-up. This is a complex variation of a reversal/withdrawal design with multiple phases and the introduction of the phases in that order was based on the authors’ hypothesis that each would increase effectiveness in that order. During the token phase, tokens were administered to children who were following the class rules at different stages during a two-hour period and could later be exchanged for back-up reinforcers. The implementation of the token reinforcement programme produced decreases in disruptive behaviour beyond what had been achieved in preceding phases (O’Leary et al., 1969). Ayllon and Roberts (1974) further evaluated token economies in mainstream classes, by implementing one with a fifth-grade class. Tokens were administered in the form of points which could be traded for a range of activities including access to a game room and extra recess time. Points were administered during reading class for reaching a certain performance criterion on a worksheet. Mean percentage of disruption across the observed students decreased when the token economy was in place compared to baseline phases, as demonstrated via an ABAB design. Not just confined to the 1960s and 70s, more recent work has continued to provide evidence for token economies in the mainstream classroom. As recently as 2020, work has been published on the efficacy of token economies as a tool in classroom management. Dejager et al., (2020) evaluated a token economy with two first-grade classes (ages ranged from 6-7 years), comparing it to a response-cost condition (five tokens made available for each student at the start of the lesson but were removed contingent on disruptive behaviour), and a combination condition where the token economy was combined with a response-cost procedure (i.e., tokens could be both awarded and subsequently removed). Their findings suggested that the token economy and combination conditions were more effective than response-cost alone, with the token economy alone being the most effective in reducing problem behaviour and increasing academically engaged behaviour across both first-grade classes. This new evidence is a welcome addition to the knowledge base, given a previous review on token economies by Maggin et al. (2011) could not uncover enough high-quality evidence to label the intervention evidence-based. This was mainly due to studies not maintaining rigorous design standards (Kratochwill et al., 2010). An example of poor rigour in design, is in an early token economy evaluation by O'Leary and Becker (1967), where an AB design was used with no withdrawal phase, thus failing to demonstrate experimental control.

Group Contingency Interventions in the Classroom Management Literature

Group contingency interventions have been deemed an evidence-based practice for targeting challenging classroom behaviour by Maggin et al. (2012; 2017) who conducted a systematic review and a subsequent updated review in recent times. Group contingencies involve the application of operant techniques in group contexts, and can be categorised as dependent, independent or interdependent (Litow & Pumroy, 1975). A dependent group contingency is established when the outcome for the group depends on the performance/behaviour of one group member or sub-group within the group. For example, making holding a class party contingent on John completing all his homework. An independent group contingency is established when a common contingency is in place for every member of the group, but individual performance determines the consequence. For example, students completing 10 problems on a worksheet by 12pm gain 5 min free time.

An interdependent group contingency is established when a certain performance criterion is in place for the group and it is the groups' collective performance which determines success. For example, if the scores on a class English test are on average 80% or greater, the whole class can take part in a class party. Maggin et al's (2012; 2017) reviews have identified the interdependent group contingency as the most common in the behaviour management literature. Group contingencies have been applied with students from kindergarten to tenth grade. They have been effective with both whole class groups and individual students within class groups and the research has mainly focused on late elementary and early middle school students (approximately equivalent to between 5th class and 1st year in Irish school settings; Maggin et al., 2012; 2017). The idea of a group contingency was first touched upon in the literature by Bushell et al. (1968), whose study was discussed earlier. However, according to Maggin et al. (2012), the earliest description of a group contingency identified in the literature was an application of the Good Behaviour Game (GBG) by Barrish et al., (1969). The following sections will discuss this classic application of a group contingency in further detail, outlining the populations with which it has been carried out and the recent developments in its application.

The Good Behaviour Game. When beginning this review of the GBG it seems important to outline the first GBG study. Barrish et al. (1969) introduced the GBG; a game to be used as a classroom management intervention targeting the occurrence of disruptive behaviours. A fourth-grade class (approximate age 9-10 years) were recruited for participation based on their high rates of disruptive behaviour, specifically out-of-seat and talking-out behaviours. Following baseline data collection during math and reading periods where high rates of these behaviours occurred, the GBG was introduced to the class during the math period. Concurrently, baseline data continued to be collected during the reading period. The teacher divided the class group into two teams and outlined the rules of the game, specifying that teams would be eligible for certain privileges upon winning the game. If a team member broke one of the outlined class rules, then their team were given a mark on the chalkboard. These marks were essentially fouls, signifying to the team that a member had broken a class rule. Teams remaining below five marks at the end of the game were eligible for the privileges, such as wearing victory tags and lining up first for lunch. If both teams surpassed the five-mark criterion, then the team with the fewest marks won the game. An ABAB reversal/withdrawal design was implemented in the math period and a concurrent AB design with a staggered baseline was implemented in the reading period. In essence, the design was a reversal/withdrawal design, with elements of a multiple baseline design incorporated to demonstrate further experimental control. The GBG produced stable and reliable changes in out-of-seat and talking-out behaviour, such that when the game was in place, both behaviours reduced substantially and when the game was briefly withdrawn, both behaviours increased immediately. This seminal study has been cited over 1300 times, and 59 times in 2020 alone (according to Google Scholar on 03/09/2020), demonstrating its sustained influence on modern research and discourse in classroom management.

Since the initial application of the GBG in 1969, it has been implemented several times across mainstream primary school populations (e.g., Donaldson et al., 2011; Medland & Stachnik, 1972; Nolan et al., 2014), mainstream secondary school populations (e.g., Kleinman & Saigh, 2011; Mitchell et al., 2015) and special education/resource settings (e.g., Flower, McKenna, Muething et al., 2014). It has been deemed very effective in targeting prosocial behaviour, problem behaviour and challenging behaviours in the classroom (Bowman-Perrott et al., 2016; Flower, Mckenna, Bunuan et al., 2014) and continues to be implemented and evaluated in the literature today (e.g., Dadakhodjaeva et al., 2019; Ford et al., 2020). Researchers have looked towards evaluating specific modifications and procedural variations of the GBG, attempting to identify the various ways it may be applied while sustaining effectiveness. These modifications have involved modifying the steps of the game (e.g., Ford et al., 2020) or modifying the conditions under which it is applied. For example, in one of the most recent evaluations, Dadakhodjaeva et al. (2019) asked if the GBG's positive effects on behaviour could be sustained if the game was played at a reduced frequency rather than daily. Using a multiple baseline across participants design, the authors demonstrated that the GBG was first effective in targeting disruptive behaviour and academically engaged behaviour when played daily in three kindergarten classrooms. During the following phase, the GBG was played at a reduced frequency. It was played on random days during the week, but data were collected on behaviour every day. The improvements in behaviour were largely sustained across this 'reduced frequency' phase, even on days where the game was not in place.

Donaldson et al. (2015) evaluated if the GBG maintained effectiveness in activities immediately before the GBG and immediately following GBG cessation across five kindergarten classes. The GBG only impacted behaviour during activities where it was played and not during activities prior to the game or immediately after the game. This is a useful point of note and points to importance of potentially modifying the game to successfully fade the intervention. In further novel evaluations, Donaldson and colleagues have examined the impact of the GBG on individual behaviour (Donaldson et al., 2017), demonstrating its effectiveness with individual students within the group context. It has also been evaluated with various implementers, such as the experimenter, and students in the class (Donaldson et al., 2018). Donaldson et al. (2018) found that across three early elementary school classes, the GBG was effective whether the teacher, experimenter or a student was implementing the game.

In another examination of a procedural variation of the GBG, Ford et al. (2020) looked at whether a ‘no-team’ version of the GBG could be effective in secondary school classrooms. Students across three classes were exposed to the game, however rather than dividing the class into distinct teams, the teacher considered the whole class as one team. This meant the GBG was still implemented as an interdependent group contingency, but behaviour may have been easier to monitor for the teacher than during a version where multiple teams were spread across the classroom. Increases in engagement and decreases in disruptive behaviour were observed when the no-team GBG was in place, indicating a potentially useful variation for use in secondary school settings which slightly reduces the workload associated with implementation.

It is evident that evaluation of procedural variations during the GBG has dominated the literature around the game’s application in recent years, with studies attempting to identify optimal conditions under which it is effective, but also whether it is effective when optimal conditions are not possible. The GBG and corresponding research on the game is here to stay, despite its foundations being laid over 50 years ago.

Positive Modifications of the Good Behaviour Game. The GBG is based on differential reinforcement of low rates of behaviour (DRL). This is described by Cooper et al. (2007, p.480) as “reinforcement [being] delivered at the end of an instructional or treatment session if during the entire session the target behaviour occurred at a number equal to or below a predetermined criterion”. However, the delivery of marks may be considered positive punishment and certainly draws teacher attention to negative and disruptive behaviour throughout the intervention. This may be viewed as problematic when school behaviour policies aim to foster positive behaviour. Irish recommendations from Tusla (the Child and Family Agency), state that promoting good behaviour should be the main goal of a school’s behaviour code (National Educational Welfare Board, 2008). This does not mean that punishment has or will be completely abandoned in Irish classroom settings and the National Educational Welfare Board recognise that sanctions are necessary as part of a positive school behaviour code. Nonetheless, positive approaches should be the initial course of action and be the focus for the teacher for everyday classroom management. Therefore, choosing to use the GBG where positive alternatives exist may not be the preferred course of action for teachers.

Positive modifications of the GBG have existed in the literature for some time. Just four years after the original publication by Barrish et al. (1969), Robertshaw and Hiebert (1973) evaluated the ‘Astronaut Game’, which involved division of a first-grade class into four teams. Class rules outlined ‘good astronaut behaviour’ for which teams could earn tokens. Tokens for each team were accumulated at the end of the day and teams who had earned the most tokens moved their spaceship one step closer to the moon (on a chart) and were given the first choice of free time activities for the remainder of class time. Although this intervention was not referred to as a version of the GBG, the same methods were applied, just maintaining a positive focus, and many recent studies have noted its similarity to a positive version of the GBG (e.g., Tingstrom et al., 2006). The game was effective in the reduction of inattentive behaviour for one target individual and increased worksheet completion for the whole class, however an AB design was used with no withdrawal phase, limiting conclusions which can be drawn. Eight years later, Fishbein and Wasik (1981) published a similar study with a fourth-grade class in a school library setting, this time referring to the positive iteration as the GBG. The librarian would award points to teams following library rules at different stages during a library session. Teams earning three out of a possible four points by the end of the session were eligible to take part in a special activity. When the intervention was in place, on-task behaviour increased, and disruption decreased relative to baseline/withdrawal phases. Patrick et al. (1998) later demonstrated a positive intervention’s effectiveness across three physical education classes, labelling it as a modification of the GBG which would hold students accountable for appropriate behaviour during a physical education class. The authors successfully targeted appropriate and inappropriate social behaviours and modified the game to incorporate the rewarding and removal of points contingent on behaviour. In recent times, these positive interdependent group contingency interventions based on the GBG, have been given the name the ‘Caught Being Good Game’ (CBGG; Wahl et al., 2016; Wright & McCurdy, 2012). The core difference between the GBG and the CBGG is in the provision of penalties/fouls versus points, respectively. The GBG, as previously outlined, involves positive punishment in the form of marks or fouls given to teams for breaking class rules or engaging in disruptive behaviour. Penalties are administered by the teacher immediately following a student’s engagement in a targeted disruptive behaviour, and the subsequent decrease in disruptive behaviour means that positive punishment has occurred. It is also based on DRL, with student teams being rewarded when penalties/fouls remain low (below a set threshold). In the case of the GBG, low rates of penalties accrued by teams, results in a team reward, serving as a positive reinforcer for low rates of the target behaviour. The CBGG, while similar, maintains a more positive focus throughout. In the CBGG, the core elements of the game are the same as the GBG in that students are divided into teams and given a set of class rules. Instead of marks for breaking rules, students are awarded points for displaying classroom appropriate behaviour. Therefore, while the GBG involves a combination of positive punishment and positive reinforcement, the CBGG relies solely on positive reinforcement to encourage appropriate behaviour.

Recent evaluations of positive versions of the GBG and the CBGG have had a variety of different focuses. Wright and McCurdy (2012) and Wahl et al. (2016) aimed to directly compare the CBGG to the traditional GBG. Wright and McCurdy (2012) used an ABAC design to demonstrate improvements in disruption and on-task behaviour across a kindergarten and a fourth-grade class. In this study, the GBG and the CBGG produced similar improvements in behaviour. Wahl et al. (2016) demonstrated that the CBGG was just as effective as the GBG in targeting engagement and disruption across four classroom settings (kindergarten x2, first/second grade mixed and second grade). Other studies have evaluated the incorporation of technology with a positive version of the GBG (Ford, 2017; Lynne et al., 2017). Lynne et al., (2017) recruited two fourth-grade classes and one first-grade class displaying high rates of disruption and low rates of academic engagement at baseline. Class Dojo, an interactive platform which allows teachers to award and display individual or team student points on an interactive whiteboard (ClassDojo, 2019), was used in the provision of points. The game was effective in targeting disruption and academic engagement across all three classes, as demonstrated using an ABAB design. Ford (2017) conducted a very similar study, recruiting four adolescent student populations attending high school. The positive GBG with Class Dojo was also effective here in targeting disruption and academic engagement. It is evident that the 'positive GBG' has been referred to simply as the GBG in some studies, with the methods section specifying that a positive modification was implemented (e.g., Ford, 2017; Groves & Austin, 2017; Lynne et al., 2017). In another study it has been referred to as the GBG-reinforcement (Tanol et al., 2010), and in more studies it is termed the CBGG (Wahl et al., 2016; Wright & McCurdy, 2012). A review of the literature should attempt to identify all positive implementations of the GBG, whether labelled the CBGG or otherwise, as at present, it can be difficult to identify these incidences given inconsistencies in language use around the game. Bowman-Perrott et al. (2016) went some way towards this in their review of GBG literature, by examining the potential effects of modifying the GBG as a moderator to its effectiveness. Rather than differentiating between 'positive' and 'negative' versions of the GBG in their moderator analysis, they specify versions as 'modified' or 'not modified', without distinguishing what types of modifications were present. Given the various terminology used to describe the positive version of the GBG and its evident surge in prevalence in the literature, a systematic review which specifies more detail than that by Bowman-Perrott et al. (2016) would be useful.

Game-based Classroom Management

It is evident from a recent review on behavioural class-wide interventions (Chaffee et al., 2017), that the GBG, both classic and modified, is one of the most common game-based interventions implemented in the behavioural classroom management literature. Other positive game-based interventions do exist however, and they provide further potential solutions to discipline issues faced in classrooms. For example, the Quiet Classroom Game (Radley et al., 2016) is a recently developed game which utilised an interdependent group contingency and iPad app to target noise level in three first-grade classrooms. The teachers were prompted every 2 min to check the classes' noise level on the iPad app, and if the class were below a certain criterion of decibel level, they received a point in the form of a smiley face on a chart. If the class obtained five of a possible seven smiley faces by the end of the game, they received a tangible reward. Not only was the Quiet Classroom Game effective in targeting decibel level, but also disruptive behaviour and academic engagement. Another intervention used in classroom management which is often described as a 'game' is Class Wide Function-related Intervention Teams (CW-FIT; Conklin et al., 2017; Wills et al., 2009). There are four key components in the CW-FIT intervention: 1) teaching desirable social skills to students, 2) minimising or totally eliminating social reinforcement for undesirable behaviour, 3) rewarding students by giving out points to groups and/or individuals for display of appropriate behaviour/skills, and 4) incorporation of a self or peer management component for students particularly at risk (Wills et al., 2009). A recent application in which elements of the CW-FIT were described as a 'game' was demonstrated by Conklin et al., (2017). One kindergarten, one second-grade and two seventh-grade class groups took part in the CW-FIT programme across several phases and group and individual data were collected on a range of behaviours. The game element was similar to a positive version of the GBG. Classes were divided into teams and could earn points for engaging in target skills. The CW-FIT produced increases in on-task behaviour, compliance, and hand-raising with concurrent decreases in out-of-seat and talking out behaviour. Despite the presence of many game-based interventions in the classroom management literature, the terminology around their use is varied. Gamification is a modern concept which may be a useful umbrella term to describe these game-based interventions. It will be discussed next.

Describing Game-based Classroom Management: Gamification. With GBG research spanning over 50 years and these novel game-based interventions gathering pace in recent times, it is important that the evidence is collated and discussed collectively. As part of this discussion, it is worth referencing the modern term ‘gamification’. Gamification is a term which began to garner empirical interest in 2010/2011 with definitions and corresponding articles beginning to make headway in the literature at this point. It is defined as “the use of game design elements in non-game contexts” (Deterding, Dixon et al., 2011, p.10), with game design elements referring to game interface design patterns, game design patterns and mechanics, game design principles and heuristics, game models and game design methods. These elements are described by Deterding, Dixon et al. (2011) as building blocks which are characteristic to games. Morford et al. (2014) described gamification as a behaviour analyst might conceptualise it, as “a way to engineer the real world by arranging contingencies to bring about game-playing (i.e., gamefulness) in a context where game-playing does not normally occur” (Morford et al., 2014, p.26). The authors subsequently compiled a list of six game-playing characteristics, with the aim of moving towards a functional definition of game-playing behaviour. These characteristics are presented in Table 1.2. with examples of how each of these characteristics may apply to the classic GBG. The characteristics put forward by Morford et al. can apply to incidences of gamification (i.e., game-play in non-game contexts) but also apply to games played in game contexts, such as chess. It seems reasonable to use the term gamification to describe the use of game-based interventions such as the GBG in the classroom (i.e., a non-game context), based on recent conceptualisations of the term. This is even though when the GBG was created originally in 1969, the term ‘gamification’ was not in existence.

Table 1.2.

Game-playing characteristics according to Morford et al. (2014) with links to the GBG

No.	Morford et al., (2014)	Good Behaviour Game Example
1	Player has a direct impact on game outcome/results	Student behaviour directly impacts fouls administered
2	Clear end goals/conditions	Goal is for team to remain under a foul criterion
3	Rules & barriers	Teacher outlines a set of rules which if broken results in a foul
4	A probabilistic outcome	Level/rate of misbehaviour impacts the outcome
5	Development of heuristics & strategies	Monitoring own behaviour/encouraging peers to behave well may work as strategies for success
6	Non-coerced initiation	Students may co-operate and engage with the game or not

Morford et al. (2014) argued that behaviour analytic research could benefit from what is known about game design and vice versa. The authors put forward the previously discussed game-playing characteristics (Table 1.2.), using behaviour analytic language to describe each characteristic. For example, the player's "Direct impact on the game outcome and results" refers to how a game player interacts with the game environment, eliciting consequences (i.e., reinforcement or punishment), which helps them to develop a sense of a predictable, controllable game-playing environment. "Rules and barriers" are said to describe the contingencies in place, letting the player know what they can and cannot do in order to progress. If gamification is considered using Morford et al.'s definition presented earlier, and the GBG aligns with the game-playing characteristics in Table 1.2., this serves to reiterate the fact that gamification is not new. This may seem an obvious deduction, however there are few discussions in the literature likening the GBG to an incident of gamification. De Byl (2013) conducted a study on gamifying university curriculum structure and in the author's discussion on gamification in education, the GBG and token economies are described as early and simple examples of gamification. Azoubel and Perger (2015) attempted to identify examples of gamification in the behavioural literature by searching the term in JABA, however found no records using the term, and instead used the simpler term 'game'. This demonstrates the paucity of discourse around the term gamification in behavioural game-based classroom management.

The field of gamification is developing at pace (11,200 hits for the term on Google scholar in 2020 alone as of September 2020), and has had positive implications in areas such as health (Johnson et al., 2016) and education (Sailer & Sailer, 2021). While many discussions of gamification emphasise the importance of technological elements, with some even specifying gamification as the use of video game design elements, it is evident that in its most basic form, gamification can be simple. The current thesis will maintain a broad focus on game-based interventions in classroom management, specifically those with a behavioural focus. Gamification is an important term which is now at our disposal to describe such interventions, and an in-depth discussion of where gamification meets ABA more generally is an important one for the field of ABA. Such a discussion is outside the scope of this thesis, however. The topic of gamification is raised here because to date, no piece of research has synthesised what game-based interventions exist in the classroom management literature. Gamification may be a useful term in identifying those interventions. Indeed, one aim of the current research was to identify the game-based interventions applied in the behavioural classroom management literature, specifically the review presented in Chapter 2.

The Current Thesis

Previous reviews of relevance have focused either on classroom management more generally (e.g., Chaffee et al., 2017; Korpershoek et al., 2016) or on specific games such as the GBG (e.g., Bowman-Perrott et al., 2016; Flower, McKenna, Bunuan et al., 2014). A review of game-based classroom management from a behavioural perspective is needed to compile a concise list of games that have been applied. Although the most popular game-based classroom management intervention, the GBG, has been subjected to systematic literature review (Bowman-Perrott et al., 2016), there is now evidence that researchers are thinking past this traditional format, to novel games which may produce similar effects (e.g., Radley et al., 2016). Such a review will serve as a starting point for the current research and one of the primary aims will be to collate and synthesise the research on game-based behavioural interventions for classroom management. Building upon this, the thesis also aims to evaluate the CBGG and its specific features as a positive strategy in diverse Irish mainstream contexts.

The Irish school landscape is diverse, with varying school types in terms of gender make-up (high prevalence of single-sex schools as well as mixed-sex schools) and DEIS status (a status designated for some schools to tackle educational disadvantage). Issues around disruptive behaviour in these diverse schools have been examined at an exploratory level (i.e., understanding the prevalence and effects of disruptive behaviour). For example, results from the Growing Up in Ireland (GUI) study, the largest longitudinal study on children in Ireland, has found that school engagement differs among boys and girls in Irish primary schools, with boys more likely to report dislike for their school and their teacher (McCoy et al., 2012). Students from working class backgrounds have been found to be significantly more likely to engage in truancy than those students from professional backgrounds (Darmody et al., 2008). Students in disadvantaged schools have also been found to have lower levels of behavioural engagement as reflected in lower rates of homework completion and attendance according to data from GUI (McCoy et al., 2014). Despite what we know about the issues faced in Irish schools, little work has been conducted at the classroom level in terms of intervening. This makes the Irish school system an ideal candidate for investigations around behavioural classroom interventions.

It is clear that evaluation of game-based classroom management interventions in an Irish context is a unique avenue which has not been examined before. The diversity of the Irish school system, paired with the clear need for intervention with certain groups in the system (e.g., in schools with disadvantaged status, all-boys schools), sets the stage for novel research with the potential for socially significant outcomes. Features examined will relate to potential time-saving for teachers (e.g., is immediate feedback a necessary component during the game?) and game sustainability over time (e.g., schedule thinning). These aims and their corresponding research objectives are expanded upon next.

Research Aims & Objectives

The research aims are as follows:

1. To understand and synthesise the body of literature around behavioural game-based classroom management strategies in mainstream classrooms

2. To evaluate game-based classroom management in the form of the positive intervention, the CBGG, and its features across diverse Irish classroom settings to add to the body of literature

Based on these aims, a set of research objectives were developed. The research objectives are set out as follows:

1. Develop and conduct a systematic literature review evaluating the application of game-based classroom management interventions in mainstream classrooms worldwide
2. Use this review to identify an intervention for further investigation in Irish mainstream settings
3. Evaluate a game-based intervention (specifically the CBGG identified in Chapter 2) across a number of diverse Irish classroom settings, maintaining a particular focus on specific variable features of the game

Thesis Outline

The current thesis consists of six chapters, each of which will now be outlined briefly.

Chapter 2 is a synthesis of game-based interventions applied in classroom management literature and examines their efficacy and features. This is presented in the form of a systematic review of the literature followed by a series of meta-analyses. Single-case literature examining the effects of game-based interventions on classroom behaviour were included and synthesised. Studies were evaluated in terms of the study design using the WWC design standards for guidance (WWC, 2017; 2020). Relevant data were then extracted from studies meeting design standards and effect sizes were calculated for experiments within these high-quality studies. A series of meta-analyses collated and combined these effect sizes, allowing for assertions to be made about the effectiveness of game-based behavioural interventions. A moderator analysis followed the meta-analysis phase of the review and investigated whether four potential moderators accounted for variance in effects across studies.

Chapters 3-5 contain four empirical studies on the efficacy of the Caught Being Good Game (CBGG) in mainstream classrooms with a focus on different features including visual feedback, schedule density and individual efficacy. The CBGG is a modified version of the previously discussed GBG, which was identified as an avenue for further research following the systematic review process. It maintains a focus on rewarding students with points for rule-following rather than giving fouls for rule-breaking, which is in line with schools moving towards more positive behavioural approaches. It was evident that research was needed on small methodological adjustments to the CBGG in diverse populations, and Chapters 3-5 set out to examine this.

Chapter 3 presents data from three mainstream secondary school classrooms. The CBGG was applied in these three first-year mathematics classrooms and the element of visual feedback was manipulated, such that sometimes students were provided with immediate visual feedback and sometimes feedback was delayed until the end of class. First-year students are approximately aged 12-13 years old and it represents the ninth year in the Irish school system, where students transition to secondary school. Teacher behaviour was also monitored across two of these classrooms to assert whether the CBGG influenced praise and reprimand rates of part-taking teachers. Social validity ratings were obtained from teachers and students.

Chapter 4 presents data from a fourth-class mainstream primary school classroom and maintains a focus on the evaluation of the CBGG with individual students as well as a class group within a single-sex boys' school context. Fourth-class students are approximately aged 9-10 years old and are in their sixth year in the Irish school system. The CBGG was evaluated using an ABAB withdrawal design and data were collected on academically engaged behaviour and three types of disruptive classroom behaviour (verbal disruption, out-of-seat behaviour, and motor disruption). A social validity rating was obtained from the part-taking teacher.

Chapter 5 presents data from a senior infant mainstream primary school classroom and evaluates the CBGG while manipulating the schedule of reinforcement during the game. Senior infant students are approximately 5-6 years old and are in their second year in the Irish school system. Group and individual data were also collected here to examine the effects of the game on both. The CBGG was evaluated using an ABAB withdrawal design, followed by three short phases where the reinforcement schedule was thinned over time. Data were collected on academically engaged behaviour and three types of disruptive behaviour. Social validity ratings were obtained from the teacher and students.

Chapter 6 is a general discussion chapter, which will integrate the findings across chapters 2-5. This chapter will address general strengths and limitations of the approaches taken in this thesis and discuss the implications of the work for practical and research purposes.

Chapter 2: Game-based Interventions for Classroom Management; A Systematic Review and Meta-analysis

Behavioural classroom management plays an important role in dealing with the issues of student disruptive behaviour and academic engagement. The importance of managing these issues appropriately at the classroom level cannot be understated given their potential effects on teacher wellbeing (e.g., Aloe et al., 2014) and various student outcomes such as aggression and academic performance (e.g., Barth et al., 2004; Pas et al., 2015). Systematic reviews in the realm of behavioural classroom management have held different focuses, with some authors considering behavioural classroom management strategies in general (Chaffee et al., 2017), group contingencies (Maggin et al., 2012; 2017) and token economies (Maggin et al., 2011). These reviews are crucially important to the field, given their focus on stringent methods of evidence synthesis, providing an overview of what works, for whom. This is particularly important in classroom management research, given the need for educators to have access to evidence-based practices to inform their approach to behaviour issues. Game-based interventions, like those discussed in Chapter 1 (e.g., the GBG, the CBGG), are a potential solution to classroom management issues and some such interventions, such as the GBG, have already been subject to systematic review (e.g., Bowman-Perrott et al., 2016).

Bowman-Perrott et al. (2016), conducted a systematic review and meta-analysis of the GBG and its variants. A large effect size was uncovered for the effects of the GBG on disruptive and desirable behaviours, confirming the efficacy of the long-standing intervention. This review dealt only with the GBG and its modifications and did not consider other game-based interventions. Maggin et al. (2012) carried out an initial review of the single-case literature on group contingency interventions for challenging classroom behaviour. Within the twenty-seven studies included in the review providing support for the evidence of group contingencies, many provided evidence for game-based interventions such as the GBG. Given that group contingencies, particularly interdependent group contingencies, can often incorporate team-based activity, it is not surprising that they are often incorporated in game-based interventions. Maggin et al. (2012) concluded that group contingencies could be labelled evidence-based for classroom management practices. This evidence was built upon in a 2017 update to the 2012 review (Maggin et al., 2017), with additional, more recently published game-based interventions being incorporated (e.g., Radley et al., 2016). Analysing game-based classroom interventions, independent of general group contingency interventions is an important next step, given they represent an important sub-section in the group contingency literature. There is also the potential to identify game-based interventions which are not grounded in group contingencies. Game-based interventions have been researched extensively and their enduring popularity (as evidenced in Chapter 1) present an opportunity for a unique synthesis of novel interventions.

Defining Game-based Interventions in Classroom Management

Lastrapes (2016) compiled a short narrative review on the use of games in classroom management, discussing the GBG, the CBGG, Tootling and the Teacher-Student Learning Game as behavioural interventions. Evidence for the four interventions are presented, with the author noting that the interactive and motivational elements of game-based interventions make them promising for widespread use in classroom management. However, the author does not outline why the four aforementioned interventions were chosen for discussion and it is unclear whether the 'Tootling' intervention is even considered as a game in the literature. To build upon this narrative review, a systematic approach to reviewing the literature is necessary, with concise inclusion criteria and clarity around what constitutes a 'game' or 'gamified intervention'.

In conducting a review on game-based interventions for classroom management, it is important to consider what constitutes a game-based intervention. Gamification, the modern term considered in Chapter 1, is an important concept in the current context, however its definition, 'the use of game design elements in non-game contexts' (Deterding, Dixon et al., 2011, p.9), isn't enough to ensure consistency in identification of papers to review. The definition is potentially ambiguous where multiple reviewers are involved due to the lack of clarity around game-design elements and which combination should be present in an incident of gamification. 'Game design elements' were put forward by Deterding, Dixon et al. (2011) in their seminal paper on the topic, however the authors specify that these elements are characteristic to games rather than a strict set of components. Deterding, Dixon et al. also specify that there is a high degree of subjectivity in the identification of incidences of gamification.

To date, there have been relatively few systematic reviews encompassing gamification as a core concept from which to take guidance. In one example, Johnson et al. (2016) examined how gamification affects health and wellbeing. Rather than compiling a concise definition of the term as to what constitutes gamification, they simply stated in the inclusion criteria that gamification must be mentioned and described as the subject of research to be included in the review. Any paper included in the review also had to clearly describe the elements of gamification involved in the intervention. An issue with using such inclusion criteria in the current review, is that to examine how games have been used in classroom settings, we must look to classic studies such as that by Barrish et al. (1969). As well as the issue of ambiguity discussed earlier, gamification, being a new term, presents further problems in adopting it for use in the review. Although gamification has been around for decades, the term itself only gathered pace in 2010. Therefore, many older studies which do not state that gamification is their main aim, are clearly conducting their research on the effects of a "gamified" or "game-based" behavioural intervention in the classroom, such that game-design elements are being used in non-game contexts.

Gamification often attracts the misconception that sophisticated technology or digital elements must be included for gamification to take place. Many examples of gamification in the literature incorporate digital elements, even recent applications of the GBG (e.g., Ford, 2017; Lynne et al., 2017). However, Deterding, Dixon et al. (2011) make clear that gamification should not be limited to digital technology. Taking all of this into account, the definition for the current review, is that the researchers, teacher or any other stakeholder delivering an intervention in the classroom, described the intervention as a 'game', an incident of 'gamification' or a 'gamified intervention' when describing it to students. This is similar to Johnson and colleagues' (2016) method, however the inclusion of the simpler term, 'game', will allow for a greater number of relevant studies to be uncovered.

In laying the foundations for a review on game-based behavioural classroom management interventions, it is important to explicitly define what will be considered a 'game-based intervention', however, it is equally important to consider what does *not* constitute a game-based intervention. Systematic literature reviews require unambiguous language in defining inclusion and exclusion criteria to ensure that they are replicable and to ensure consistency across multiple reviewers who may be involved. In considering this, serious games must be discussed. Serious games differ from gamification in that in many definitions they have a technological element. For example, Girard et al. (2013) used a definition by Marsh, (2011) in their recent meta-analysis on serious games and this definition clearly stated that a serious game included a digital, virtual or simulated element. Game-based interventions such as the GBG may incorporate a technological element (e.g., Class Dojo), however most iterations are low-technology, with points administered on the board (e.g., Barrish et al., 1969). In addition, Deterding, Dixon et al. (2011) stated that while gamification may incorporate some but not all game design elements, serious games are fully-fledged games, not designed for entertainment purposes (e.g., to teach a skill).

While the GBG is a game-based intervention aimed at encouraging positive behaviour during regular instructional periods like a maths class, a serious game may involve use of digital technology to teach students to behave in a positive manner, and would not be played during regular instructional periods. It is therefore not within the remit of this review to incorporate serious games.

Systematic Review of Single-case Research Designs

The current review will focus solely on single-case design evaluations of game-based interventions in the classroom. Although previous research has evaluated game-based interventions using alternative methods, such as randomised controlled trials (e.g., Ashworth et al., 2020), single-case research designs maintain unique strengths for evaluating such interventions. Single-case research designs tend to focus on overt, observable behaviour rather than self, teacher or peer report and based on scoping searches in the area, it was supposed that a majority of game-based behavioural interventions applied in the literature were applied using a single-case approach. A previous review by Smith et al., (2019) collated studies on the GBG which used randomised controlled trial (RCT) designs. This review identified seven studies for inclusion in a meta-analysis and reported small effect sizes across various student outcomes. This contrasts with recent meta-analytic findings on single-case research applications of the GBG which have reported larger effect sizes across a larger number of studies (Bowman-Perrott et al., 2016). However, it is evident that many target behaviours in RCTs were measured using teacher report measures of student behaviours at various time points across a study. Single-case research designs are inherently different in that behaviour is measured by trained observers during every iteration of the intervention. This difference in outcome measures is one potential reason for the difference between effect sizes reported in reviews on RCTs (i.e., Smith et al., 2019) versus reviews on single-case research designs. Chapter 1 outlined the usefulness of single-case research designs in the evaluation of behavioural classroom management interventions, and these research designs will be the only type considered in the current review. The Institute of Education Science’s (IES) What Works Clearinghouse (WWC) division have put forward sets of guidelines which may be used in the evaluation of the design quality of single-case research designs. These guidelines were first developed by Kratochwill et al., (2010) for the WWC and up to date versions of the standards have been published by the WWC in recent years (WWC, 2017; 2020). In 2017, version 4 of the WWC guidelines was published and these guidelines outline four key stages in evaluating a study’s design quality. These are:

1) The independent variable is systematically manipulated such that the researcher determines when and how the conditions change;

2) Interobserver agreement (IOA) data are collected periodically during the study, at least once per phase, at least 20% of the time in each study condition and the agreement between the observers should be over 80%;

3) There should be at least three attempts to demonstrate intervention effects over time;

4) There should be a minimum of 3-4 data points per phase for the study to meet WWC standards with reservations and at least 5 data points per phase for the study to meet WWC standards fully (WWC, 2017).

Therefore, there are three potential ratings a study can be given: *Meets WWC Standards*, *Meets WWC Standards with Reservations* or *Does Not Meet WWC Standards*. In version 4 of the guidelines (WWC, 2017), standards for study evidence are also presented and guidelines for assessing the evidence of a study visually are outlined. An updated version of these guidelines was published in 2020 (version 4.1.; WWC, 2020), with the additional design criteria that the data should be presented in graphical or tabular format. The visual analysis guidelines are not outlined in this updated version, and authors are encouraged to compute a design-comparable effect size to interpret the magnitude of the study findings where appropriate. The current review was prepared across 2018 and 2019 and therefore version 4 of the standards was adhered to when conducting study design evaluations. However, no study without graphed data was included, therefore the additional criterion put forward in version 4.1. was met. The review was underway when version 4.1. was published (WWC, 2020), however this version was consulted before the evidence evaluation took place. Therefore, a design-comparable effect size was computed for all studies meeting the design standards with or without reservations.

Versions of the WWC design standards have been applied in many reviews of single-case research designs in recent times. For example, Maggin et al. (2012; 2017) used the WWC standards in their syntheses of group contingency interventions and Bowman-Perrott et al. (2016) used the standards when synthesising GBG research. There does appear to be some variability with exactly how authors implement these standards. For example, Bowman-Perrott et al. specified that IOA should be at least 80% and collected during at least 20% of baseline and/or intervention phases but did not specify that it should have been collected at least once per phase. Maggin et al. (2017) similarly specified the

IOA should be collected at least 20% of the time overall with at least 80% agreement. There is also some variability in how authors have used the design standards to inform the remainder of their review. For example, Bowman-Perrott et al. calculated effect sizes for all studies included in the review after the full-text screening phase, as long as they met two basic design standards criteria; that they could demonstrate experimental control and that they had at least three data points per phase. Maggin et al. (2017) only calculated effect size data for studies meeting the design standards with or without reservations. To ensure only the highest quality of studies are included in the current synthesis, only studies meeting the WWC design standards with or without reservations will proceed to the next phases of the review (i.e., study characteristic coding and effect size calculation).

Overall Purpose of Synthesis

Researchers continually strive to produce new and effective game-based behavioural interventions for the classroom (e.g., Radley et al., 2016) as well as modifying and modernising classic games (e.g., Wahl et al., 2016). A review on the use of game-based behavioural interventions in the mainstream classroom has not yet been conducted therefore that is what the review in this Chapter seeks to address. Although there have been many applications of game-based interventions in special education and resource classrooms (e.g., Flower, McKenna, Muething et al., 2014; Groves & Austin, 2017), this review focuses on mainstream classrooms because they tend to be inherently different from special education classrooms. For instance, classes tend to be larger and teachers may not have access to the same supports for classroom management and challenging behaviour as those teachers in special education classrooms. It is therefore important to consider these settings separately, rather than combining different settings in one review. Outcomes of this review may inform future research on reviewing the game-based intervention literature in special education classrooms. Group contingencies, many of which are game-based, have been reviewed in recent times, however the important intersection of group contingencies and game-based interventions has not been. Additionally, game-based interventions which do not incorporate a group contingency may exist and it is important to identify these. Although it is likely that many of the papers identified during this review will include evaluations of the GBG, language on the use of the positive version of the GBG has been ambiguous in recent years, as outlined in Chapter 1. This review will therefore clearly identify positive iterations of the GBG and examine the language used to describe these interventions. Game-based interventions have the potential to enhance classroom experiences through enjoyment for students, while being low-tech and easy to implement for teachers. This review will assist in generating a concise synthesis of what has been done, as well as identifying gaps in the literature for future research.

Research Questions

A number of research questions are proposed:

RQ1: What game-based classroom management interventions exist in the mainstream classroom management literature?

RQ2: How are these game-based classroom management interventions being applied in the classroom? i.e., Who is responsible for implementing the game? What are the specific rules/objectives? What types of prizes/rewards are being administered?

RQ3: What specific behavioural outcomes have game-based classroom management interventions targeted amongst mainstream school students?

RQ4: Have game-based classroom management interventions been effective in targeting classroom-based behaviour?

RQ5: What are the effects of specified moderators on the effectiveness of game-based classroom management interventions on behavioural outcomes?

Method

Search Strategy

A search strategy was devised by the student researcher and this strategy was reviewed by the primary research supervisor and the Dublin City University subject librarian before being finalised. The aim of the search was to uncover studies which reported on game-based classroom management interventions applied in mainstream classroom settings. Single-case research designs were the design type considered for this review. The search terms were therefore devised under four main categories: 1) gamification and game-related terms, 2) behaviour terms (e.g., behavio#r, “talking-out”), 3) setting terms (e.g., class*, school, “first-grade”) and 4) design terms (e.g., “single-subject”, “multiple baseline”). Five databases were searched: PsycInfo, Education Research Complete, ERIC, Web of Science and Scopus. A comprehensive description of the exact search terms used and strategies for individual databases can be seen in Appendix A. Following searches across the five databases, 2451 records were exported to data management software Refworks for the first stage of screening. Figure 2.1. provides an overview of the search, screening, and identification process, based on the PRISMA flow diagram (Moher et al., 2009).

Screening

Duplicate Removal

The first stage of screening involved removal of duplicate records (i.e., records which had been uncovered in more than one of the databases). The Refworks tool for removing ‘exact duplicates’ and ‘close duplicates’ was used, with all records checked by the student researcher before final removal. The records were then checked alphabetically, first by author and then by title to ensure all duplicates were removed. After removal of duplicates (n = 480), 1971 records were retained for title and abstract screening. During the screening phase, two more duplicates were identified and subsequently removed, bringing the total number of records to be screened to 1969 and total number of duplicates to 482.

Title and Abstract Screening

Title and abstract screening was conducted independently by two reviewers. The two reviewers were the student researcher and Dr. Ronda Barron, a lecturer and researcher holding a MSc. in ABA and PhD in Psychology. From this point forward, the ‘first reviewer’ will refer to the student researcher and ‘the second reviewer’ will refer to Dr. Barron, unless otherwise specified. All records were copied from Refworks to online citation screening platform ‘Abstrackr’ (Wallace et al., 2012). This platform allows two or more screeners to screen abstracts and deal with conflicts in decision-making. Inclusion and exclusion criteria were applied at this stage however, screeners were not required to provide a reason for exclusion. Inclusion and exclusion criteria are presented in Table 2.1. and reflect the aims of the review. The first reviewer tagged any relevant reviews at this stage. Reviews were not included as studies in the current review, however the reference lists of reviews were checked for potentially relevant records.

Upon screening the 1969 records, the two reviewers made conflicting decisions on 10.8% of occasions. The reviewers subsequently met to discuss all discrepancies and come to a final decision. A third reviewer (research supervisor, Dr. Sinéad Smyth) assisted on decisions upon which the two reviewers could not agree. In total, 1803 records were excluded at this stage. One-hundred and sixty-six records were retained for full-text screening.

Table 2.1.

Criteria for Studies to be Included in the Review

Criteria	Example of Included	Example of Excluded
Year: Study must be conducted after 1960	Study conducted in 1960 or afterwards	Study conducted in 1959 or before
Language: Study should be published in English	Study is written in the English language	Study written in any language other than English
Publication type: Study must be an ***empirical investigation*** Any empirical study identified in searches is considered eligible except for masters theses (doctoral theses are eligible)	Study reports on an empirical investigation of an intervention.	Reviews, books, conference proceedings, masters theses, instructional pieces (e.g. an instruction manual on how to use the Good Behaviour Game).
Independent variable: Intervention is described explicitly as a game/gamification in methods section.	Study describes the application of the Good Behaviour Game intervention where the intervention is described in the methods section as a 'game' or as the 'good behaviour game'.	A study which describes an intervention which may appear to mimic a game, but is not explicitly described as a 'game'.
Population: The intervention must have been applied with a group of students enrolled in mainstream K-12 education (or its equivalent across cultures)	Study investigates a game-based intervention with a group of 4^{th} grade students attending a general education classroom.	Study investigates a game-based intervention with a group of students with ASD attending a special school.

Criteria	Example of Included	Example of Excluded
Setting: The intervention must have been implemented in a classroom setting	Study investigates an eligible intervention in a 4th grade classroom setting.	Study investigates an eligible intervention with mainstream students, but it is implemented in the school library.
Dependent variable: The DV must be an observable classroom-based behaviour related to classroom management issues e.g., disruptive behaviour, on-task behaviour, engagement, out of seat behaviour, attendance/punctuality, social behaviour	Study investigates the impact of the Good Behaviour game on the on-task and disruptive behaviours of a 4th grade mainstream maths class.	Study investigates whether the good behaviour game leads to increased accuracy on a Maths worksheet (i.e., the behaviour is related to academic achievement rather than academic engagement).
Design: Study should be a single-case research design (i.e. reversal, changing criterion, multiple baseline, multi-element or any variation of these design types)	Study investigates the impact of the Good behaviour game on disruptive behaviour of 4th grade students. A reversal design is used to assess intervention effectiveness.	Study is a randomised controlled trial evaluating good behaviour game effectiveness.

Full-text Screening

Full-text screening was conducted independently by the first and second reviewers and took the same inclusion/exclusion criteria into account as during the title/abstract screening phase (see Table 2.1.). Initially, full texts of 159 of these studies were obtained through the university library system or by contacting authors of these articles. Seven studies could not be retrieved. Reviewers recorded the reason for exclusion during this phase and exclusion criteria were applied in a systematic order, as presented in Table 2.1. This stage of review was managed in the Zotero referencing platform as a more stable platform than Refworks. Zotero is also the referencing platform recommended by the DCU library.

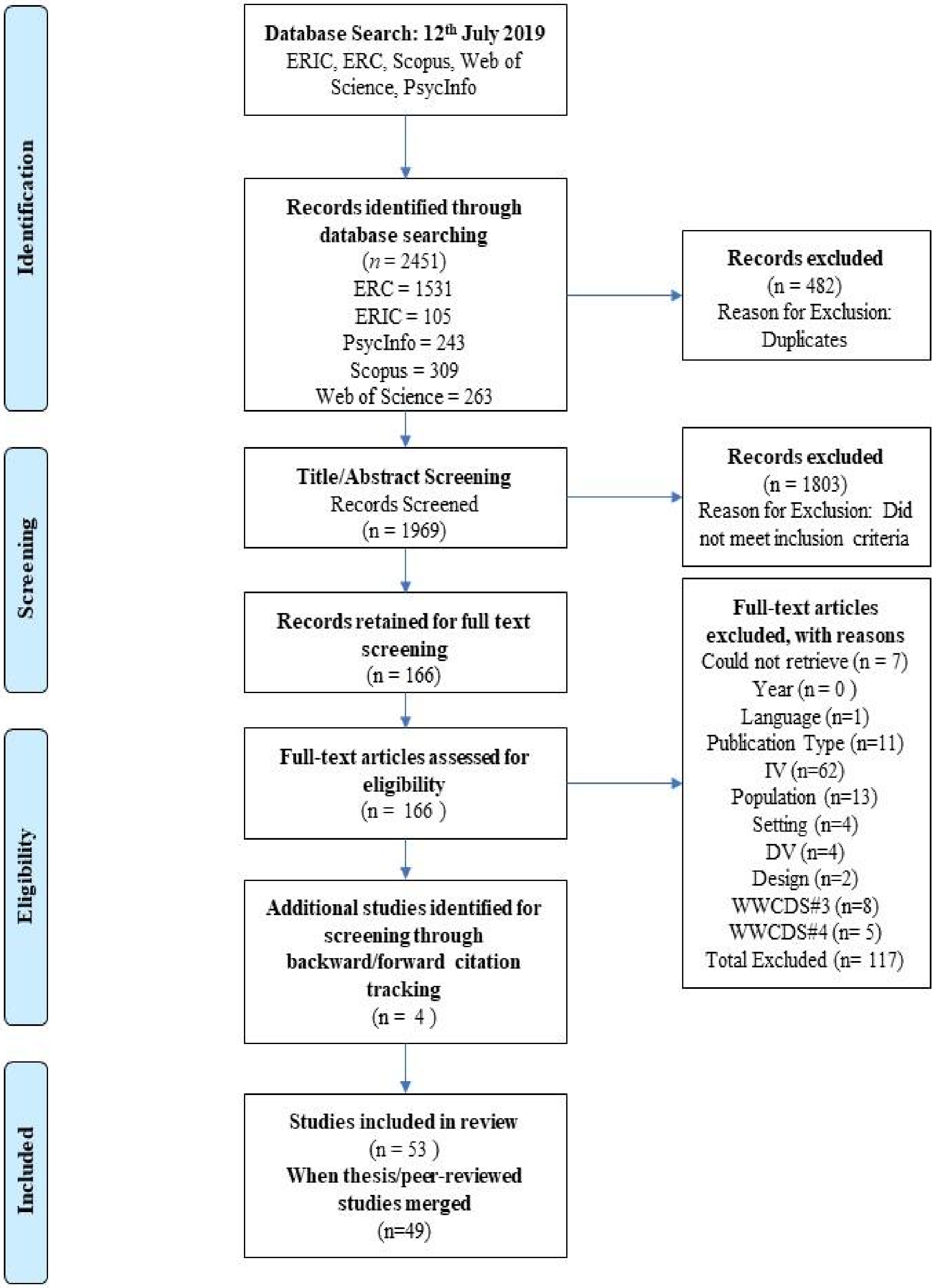
Two additional inclusion criteria were added at this stage of screening in order to streamline the process and to ensure the next phase would include studies of the highest possible quality. These two criteria were two of the What Works Clearinghouse (WWC) design standards (Standard 3 and Standard 4; WWC, 2017), which are used to assess the design quality of single-case research designs. Design standard 3 (DS#3) stated that studies must involve three attempts to demonstrate intervention effectiveness (e.g., an ABAB design, or a multiple baseline design with at least three concurrent baselines). Design standard 4 (DS#4) stated that there must be at least three data points per phase for a study to meet WWC standards with reservations. These two standards were applied at this stage because reviewers could apply them quickly by scanning the graphs included in the results section of included studies. If there were multiple graphs/experiments in a study, and at least one of them met these two standards, then the study progressed to the next phase of the review for further probing. The WWC design standards as applied in this review will be discussed in more detail later in the Methods section, in the "Study Design Quality Evaluation" subsection.

After this screening phase, the two reviewers again met to discuss discrepancies. Reviewers disagreed on 11.3% of the studies available for full text screening (18 out of 159). At least one of the reviewers rated another 13.8% of the studies with a 'maybe' (22 out of 159). These discrepancies were resolved during a meeting and discrepancies that could not be resolved were discussed with a third reviewer. In total, 49 studies were retained for inclusion in the review. Reasons for exclusion of 117 studies are outlined in Figure 2.1.

Upon screening a study and subsequently including it as it met the criteria, the first reviewer checked the reference list of the study (backward citation tracking). If the title and abstract indicated that a reference was potentially relevant, then the study was retrieved in full text to review. This practice identified four further studies for inclusion. Therefore, the total number of studies included was 53. Of these 53 studies, four pairs of studies were found to be the same study in thesis and peer-review format. The peer-reviewed version of the study was retained for inclusion in this case. Therefore, 49 individual studies were included in the review. Each of these studies were assigned an ID number from 1-49. Individual experiments within studies were also assigned an ID number based on the study ID. For example, the first experiment in Study 1 would be assigned the ID 101, the second experiment in Study 1 would be assigned the ID 102, and so on. An 'experiment' refers to one outcome and one participant/participant group in a withdrawal/reversal, alternating treatment or changing criterion design. For example, in a study by Ford (2017), a positive version of the GBG was applied across four classrooms using a reversal/withdrawal design. Two outcomes (academic engagement and disruptive behaviour) were monitored in each classroom. There were therefore eight individual experiments in this study (four classrooms X two outcomes). In a multiple baseline design, an experiment refers to one outcome but may include more than one participant across three or more staggered baseline phases, such that one multiple baseline design and outcome is one experiment. For example, in a study by Dadakhodjaeva et al. (2019) the GBG was applied across three classrooms using a multiple baseline across participants design. Two outcomes (academic engagement and disruptive behaviour) were monitored in each classroom and additionally, the behaviour of one target student was monitored in each classroom. This study therefore contained four experiments.

Figure 2.1.

Flow Diagram Outlining the Study Selection Process for the Review



Study Design Quality Evaluation

The design of studies included in the current review were evaluated using the WWC's four main single case research design standards outlined in version 4 of the WWC standards handbook (WWC, 2017). These guidelines provide an overview of standards which single-case research designs should meet and allow for a rating to be assigned to 1) experiments within studies and 2) the study as a whole. There are three possible ratings when examining a study under the WWC criteria: the study *Meets WWC Standards*, *Meets WWC Standards with Reservations* or *Does Not Meet WWC Standards*. A Microsoft excel file was created outlining the key design standard questions and reviewers completed one excel sheet per included study (Appendix B). This excel document is based on the WWC study review guide (WWC, 2016) and accompanying excel file on the WWC website, such that elements of the WWC study review guide were copied exactly as written to ensure close compliance with the guidelines. The results were then inputted to a master excel sheet where all results could be collated.

The WWC designs standards were applied as outlined earlier in the Introduction section to this chapter (WWC, 2017). To reiterate, the first WWC design standard (DS#1) states that the independent variable (i.e., the game-based intervention) is systematically manipulated by the researcher. The second standard (DS#2) is concerned with ensuring studies meet acceptable thresholds for collection of IOA data during the study. To meet DS#2, IOA must have been collected once for each outcome per study phase, on at least 20% of sessions overall and meet the minimum acceptable mean value of 80% agreement for each outcome variable. If a study included some experiments with outcomes meeting this standard and some which did not, the study was considered to have met the standard overall and only experiments with outcomes meeting the standard progressed to the next phase of the review.

Studies had already been evaluated under DS#3 during the full-text screening phase (studies must include three attempts to demonstrate an intervention effect), however at this stage in the design evaluation, reviewers had to decide how many and which experiments within the study met this standard. Studies had also been evaluated under DS#4 (experiments should have 3-4 data points per phase to meet standards with reservations or 5 points per phase to meet standards), but again, reviewers had to be more specific at this stage in identifying experiments within the study meeting and not meeting this standard. Also included in the design evaluation phase, was a section for the reviewer to note any potential confounds in the study (e.g., the intervention agent changed throughout the study, the method of data collection changed throughout the study). Any confounds noted were subsequently discussed by the two reviewers. If the confound was considered problematic, the study did not progress to the next phase. The nature of school based research means that the potential for confounds is prevalent, despite researcher attempts to demonstrate experimental control. Many confounds could be unpredictable and therefore a list of potential confounds was not developed prior to this review. In deciding whether a confound was problematic, the two reviewers would consider the confound in the context of the study as a whole, paying attention to the graphed data. For example, in the study by Mitchell et al. (2014), a change in teacher occurred in one of the classrooms approximately half way through the study. Both reviewers noted this was a potential confound, however on examining the data for consistency and discussing the fact that teacher changes are a regular occurrence in school settings, it was determined that this was not a problematic confound. The reviewers were finally required to outline each experiment under scrutiny within the study, and make a decision on whether it met the design standards with or without reservations or did not meet standards. An overall study rating was then assigned. In determining the overall study rating, the ratings given to each experiment within the study were considered. The highest rating given to any experiment within the study became the overall study rating. Therefore, if a study contained some experiments that did not meet the WWC design standards, but contained even one that met the standards fully, then the overall study rating was *Meets WWC Standards*. Only studies rated *Meets WWC Standards* or *Meets WWC Standards with Reservations* proceeded to be evaluated further (i.e., study characteristic coding and effect size calculation).

Study Characteristic Coding

Studies which satisfied the initial inclusion criteria and met the WWC design standards (WWC, 2017; 2020) were coded for a number of variables. The first reviewer extracted data from all included studies and a postgraduate research student in the School of Psychology extracted data from a sub-set of 10 studies (33.33% of the total amount included). The two reviewers agreed on 95.53% of occasions and this was rectified to 100% agreement after discussions. Before fully coding a study, reviewers took note of the number of experimental cases in the whole study, number of experiments meeting standards with or without reservations, number of participants cases in the study, the publication type, country, research aims and study design. Data were then extracted from the study under a number of different headings.

Student/School Characteristics

The reviewers extracted data on the students receiving the intervention and the setting in which the intervention took place. This included a) the number of students receiving the intervention, b) the gender composition of the sample, c) description of the sample (i.e., individual or group), d) ethnicity of the sample e) mean age of sample, f) age range of sample, g) grade level of sample and h) the subject ongoing while the intervention was put in place (e.g., Maths class, circle time).

Study Methods and Intervention Implementation

A range of data were extracted on the study methods and intervention implementation. These included a) target behaviour(s), b) how the target behaviour was measured, c) intervention/game name, d) intervention/game description, e) rewards available for winning, f) reward frequency, g) intervention/game implementer, h) implementer training and description if relevant, i) treatment integrity and rate of treatment integrity, j) teacher social validity and k) student social validity. The reviewers also wrote a summary of the results for each experiment within the study.

Data Analysis

Graphed Data Extraction

Graphed data were extracted from studies meeting design standards with or without reservations using the online WebPlotDigitizer software (Rohatgi, 2019). This software has been deemed reliable, valid and user-friendly for the extraction of graphed data from single-case research studies (e.g., Drevon et al., 2017; Moeyaert et al., 2016). The software has been used by other authors conducting reviews of single-case research design studies (e.g., Barton et al., 2017; Common et al., 2020; Long et al., 2019; Maggin et al., 2017; van Dijk & Gage, 2019). The first reviewer extracted data from 143 graphed experiments (84.6% of experiments meeting standards with or without reservations). In some cases, data could not be extracted because graphed data was too difficult to extract reliably (Conklin et al., 2017; Mitchem et al., 2001; Patrick et al., 1998; 19 experiments in total) and seven experiments were alternating treatment designs therefore effect sizes were not to be calculated and data was not extracted (Hartman & Gresham, 2016; Hernan et al., 2019). X-axis data were specified as on the author’s study graph, in other words, if there were breaks specified clearly on a graph where data were not collected for a number of sessions, this was accounted for in extracting data. This was not possible for one study (Fallon et al., 2018) as it was difficult to specify where the breaks in data collection occurred on the X-axis.

In order to obtain a measure of reliability, a third-year undergraduate psychology student (who was on work placement with the project supervisor) extracted graphed data from five studies (i.e., 16.67% of studies meeting the design standards with or without reservations), including 29 experiments (20.3% of the 143 experiments for which data was extracted; Dadakhodjaeva et al., 2019; Ford, 2017; Hine et al., 2015; Kleinman & Saigh, 2011; Lynne et al., 2017). All correlations between the first reviewer's data and the undergraduate student's data across the 29 experiments were greater than .99. As a second method of checking the reliability of extracted data, the first reviewer checked the mean of each phase of data extracted against the mean reported in the study (if reported). If the means differed by more than 1.5, the data were re-extracted.

Effect Size Computation

There is no single recommended effect size metric for single-case research designs and recent reviews of single-case research have adopted an approach where a number of effect sizes are calculated (e.g., Gage et al., 2020; van Dijk & Gage, 2019). Non-overlap effect sizes, such as Tau-U and Percentage of Non-Overlapping Data (PND) are often used to evaluate the magnitude of effect in single-case research designs (e.g., Bowman-Perrott et al., 2016; van Dijk & Gage, 2019; Watkins et al., 2019). These are non-parametric methods which are generally easy to use, blending well with visual analysis of data (Parker, Vannest & Davis, 2011). In their most recent set of standards (version 4.1.), WWC recommend calculating design comparable effect sizes where appropriate (WWC, 2020). These are effect sizes which can be compared with a standardized mean-difference effect size commonly computed for group-design studies and used by the WWC for group design studies (WWC, 2020). Based on the above discussions, it was decided that two effect sizes would be calculated here; one non-parametric non-overlap effect size (Tau) and one design comparable, regression based effect size (Between Case-Standardised Mean Difference [BC-SMD]; Hedges et al., 2012; Pustejovsky et al., 2014) recommended by the WWC (2020).

In calculating effect size metrics, some study and experiment level considerations had to be taken into account. Some studies evaluated the effects of two different versions of an intervention (e.g., Tanol et al., 2010; Wright & McCurdy, 2012). Where this occurred, any phase where a game-based intervention was applied was considered a 'game phase' and thereby included in the effect size calculation. For example, in the study by Tanol et al., (2010), an ABACBC design was applied across two classrooms, with A representing the baseline phase, B representing the GBG response-cost and C representing the GBG reinforcement. Although the phases reflect implementation of two different types of the GBG, the current review was concerned with game-based interventions in general and therefore the B and C phases were merged as one ‘intervention phase’ when calculating effect size. Therefore, the effect size reflects the overall effect of all versions of game-based intervention applied in the study. A similar protocol was followed where the same intervention was applied by different individuals across phases (e.g., Donaldson et al., 2018) and where a ‘training’ or ‘feedback’ phase preceded intervention implementation (e.g., Wright & McCurdy, 2012).

Non-parametric Effect Size: Tau. The Tau effect size was calculated for each phase contrast in studies meeting the WWC design standards (WWC, 2017) with or without reservations. It is important that there is a distinction made between the Tau effect size calculated here (Tarlow, 2017), and the commonly used Tau-U effect size (Parker, Vannest, Davis & Sauber., 2011). The Tau family of effect sizes are based on Kendall’s rank correlation and Parker, Vannest, Davis and Sauber (2011) put forward the $\text{Tau}_{\text{novlap}}$ and Tau-U metrics. $\text{Tau}_{\text{novlap}}$ is obtained from Kendall’s rank correlation or the Mann-Whitney U test, but Parker, Vannest, Davis and Sauber (2011) specify that minor adjustments be made to the formula if using Kendall’s rank correlation because it is not designed for dummy-code variables. The extension of the $\text{Tau}_{\text{novlap}}$ effect size, Tau-U, allows for control of undesirable baseline trend and both $\text{Tau}_{\text{novlap}}$ and Tau-U can be calculated using a free online calculator (Vannest et al., 2016). Tarlow (2017) discusses some limitations around the Tau effect sizes suggested for use by Parker, Vannest, Davis and Sauber and put forward Baseline-corrected Tau which produces more conservative Tau values. Before baseline correction, Tarlow suggests using dummy code variables to calculate Tau, because the alteration to the formula put forward by Parker, Vannest, Davis and Sauber, can lead to inflated values for Tau, falling outside of the bounds of -1 and +1. The subsequent baseline correction put forward by Tarlow is based on Theil-Sen regression methods. Given the more conservative values of Tau and subsequently Baseline-corrected Tau provided by the use of Tarlow’s (2016) Tau calculator, it was chosen for use in this review.

Tarlow’s Baseline-corrected Tau calculator (Tarlow, 2016) was used to calculate a Tau value and standard error (*SE*) for each phase contrast within an experiment. Baseline correction was implemented where the calculator suggested a significant trend was present in the direction of proposed treatment effect. A weighted average Tau value and *SE* could then be calculated for each experiment within a study by inputting the Tau value and *SE* for each phase contrast into Winpepi (Abramson, 2011) and using the meta-analysis function. A similar method was followed by Bowman-Perrott et al. (2016) in a systematic review on the GBG and by Fallon et al. (2018) in a review on direct training for treatment integrity. Specifically, this involved opening the Winpepi for desktop application, selecting '*COMPARE2 (comparison of two independent groups or sample)*' on the first screen and '*META-ANALYSIS; analysis of stratified data*' on the second screen. Then, '*Other**, of proportions or rates with SEs/CIs*' was selected under '*WHAT STATISTICS WILL BE ENTERED?*' and '*Standard error*' was specified under '*ALSO ENTER*'. One key difference between the current review and those by Bowman-Perrott et al. (2016) and Fallon et al. (2018), is that the Tarlow (2016) calculator was used here rather than the calculator put forward by Vannest et al. (2016). The reason for this decision was outlined in the previous paragraph. Tau values are interpreted using Vannest and Ninci's (2015) recommendations of .20 as a small effect, .20-.60 as a moderate effect, .60-.80 as a large effect and .80+ as a very large effect. This interpretation has been deemed appropriate given the similarity between the two measures (i.e., that proposed by Parker, Vannest, Davis & Sauber, 2011 and that proposed by Tarlow, 2017; Dillon et al., 2019).

Regression-based Effect Size: BC-SMD. The BC-SMD effect size was put forward by Hedges et al., (2012) and estimates a between-case effect size for a study, corresponding to Cohen's *d* (standardized mean-difference effect size). Pustejovsky et al. (2014) built upon Hedges et al's., (2012) recommendations and developed a BC-SMD effect size which is more flexible and accommodates data with trend present (Gage et al., 2020; Valentine et al., 2016). This effect size is appropriate given WWC's (2020) recent recommendations around the use of a design-comparable effect size in single-case research. The BC-SMD effect size can be calculated for studies where a withdrawal/reversal design was applied across at least three participants for the same outcome and studies where a multiple baseline design was applied across three participants. This requirement stems from the fact that BC-SMD involves variation across participants in the outcome (Valentine et al., 2016). This means that the BC-SMD effect size could not be calculated for all studies included in the review (e.g., alternating treatment designs, studies including reversal designs with only one participant case etc.).

In the current review, the Pustejosky et al. (2014) BC-SMD effect size was estimated using a web application powered by R, *scdhlm* (Pustejosky, 2016). The accompanying tutorial paper (Valentine et al., 2016) was used in assisting with calculations. Within the *scdhlm* application, the restricted maximum likelihood (REML) estimator was used as the estimation method. The *scdhlm* application provides a number of options when using the REML method. First, at baseline and treatment level, fixed and random effects can be specified. The application requires that **1)** a fixed effect be specified for baseline (allows the intercept across all baseline phases to differ from zero), **2)** a random effect also be specified for baseline level (assumes that the level of the outcome variable will vary across cases) and **3)** a fixed effect be specified for the treatment phase (permits the intercept to vary from the baseline level; Valentine et al. 2016). Second, the application allows the user to choose whether a random effect is specified for the treatment phase. This permits the treatment effect to vary across cases. Third, when calculating effect size for a multiple baseline design, there is an additional option to specify types of time trend at baseline and treatment phases. Valentine et al. (2016) recommended that if the user is planning to synthesise the effects for inclusion in a meta-analysis, the model specifications should be the same across studies. Therefore, in the current analysis, for all included studies, a fixed and random effect was specified at all baseline phases and a fixed and random effect was specified at all intervention phases. Additionally, time trends were specified as ‘level’ across all studies. No specific guidelines have been put forward for interpretation of BC-SMD, however in a recent review and meta-analysis, Allen et al. (2020) suggest following Shadish et al.’s (2016) guidance in interpretation: small = 0.37 to 0.98, medium = 0.98 to 1.87 and large > 1.87.

Meta-analysis

Meta-analysis involves computation of an effect size and its variance for each study included in the synthesis in order to understand the results of a study in the context of the other included studies (Borenstein et al., 2009). It also allows for the computation of an omnibus effect size across included studies. In the current review, upon calculation of effect sizes, the data were divided into two outcome categories: outcomes targeted for increase (e.g., academic engagement) and outcomes targeted for decrease (e.g., disruptive behaviour), which would be considered in separate meta-analyses. This resulted in four data sets for meta-analysis; positive behaviour Tau, disruptive behaviour Tau, positive behaviour BC-SMD and disruptive behaviour BC-SMD.

Some studies included evaluation of the same outcome across multiple class groups or individuals (e.g., multiple reversal/withdrawal designs evaluating the GBG with several individuals in the context of one class group; Donaldson et al., 2017). Where this occurred, experiments were combined using the meta-analysis function in Winpepi (fixed-effects) in the same way as effects were combined for experiments as described in the Tau subsection in the Effect Size section earlier. One effect size was used for the specific outcome going forward (Borenstein et al., 2009). In most cases where data were combined in this way, data had been collected within the same school (e.g., Wright & McCurdy, 2012) or was reported to be collected within the same school district (e.g., Radley et al., 2016). In some studies, different outcomes were evaluated across the same participants. Where this is the case, more weight will be assigned to these studies in meta-analysis and the precision of the summary effect will be overestimated (Borenstein et al., 2009). To control for this in the current analysis, outcomes were combined by calculating a synthetic effect size which accounts for the correlation between the outcomes (Borenstein et al., 2009). Outcomes were only combined if they were included within the same outcome category. In other words, behaviours targeted for increase (e.g., engagement, compliance) were combined and behaviours targeted for decrease (e.g., talking-out, out-of-seat) were combined. Given the nature of study outcomes included in the review, the outcomes combined in each category were generally topographically similar (e.g., combination of on-task behaviour and compliance). An example of where this was applied was in a study by Radley et al., (2016) where three classes took part in the study evaluating the effects of the Quiet Classroom Game on disruptive behaviour and decibel level. These two outcomes were assumed to be highly correlated. To create a synthetic effect size for these outcomes combined, the effect size was calculated as the mean of the effect size across outcomes ($\underline{Y}$) , and the variance was calculated using the formula put forward in Borenstein et al., (2009); $Var_{\underline{Y}} = 1/4(Var_{Y1} + Var_{Y2} + 2r\sqrt{Var_{Y1}}\ \sqrt{Var_{Y2}})$, where $\underline{Y}$= synthetic effect size created, Var = the variance, $Y1$= effect size from experiment 1, $Y2$= effect size from experiment 2 and r = the correlation between the two outcomes using the extracted data.

Effect sizes and their variance (i.e., SE^2) were collated in a Microsoft excel sheet and uploaded to the *metafor* package in R (Viechtbauer, 2010). This is a package providing functions for conducting meta-analyses which has been used in other recently conducted meta-analyses on single-case research designs (e.g., Gage et al., 2020; van Dijk & Gage, 2019). A random-effects model was chosen as it was assumed that the true effect size was not the same across studies, given the diversity in samples across the included studies (Borenstein et al., 2009). Studies included in this review evaluated various game-based interventions with various school-going populations and therefore a fixed-effects model may not have been appropriate. Models were estimated using the restricted maximum likelihood estimator (REML) which is recommended as an unbiased and efficient method (Viechtbauer, 2005; 2010) and has been used in reviews in similar topic areas (e.g., vanDijk & Gage, 2019). Q and I^2 statistics were calculated and subsequently used to assess heterogeneity. Cochran's Q is a test for heterogeneity which is the sum of the squared deviations of study's estimates, weighted as in the meta-analysis (Cochran, 1954; Higgins et al., 2003). I^2 provides a measure of inconsistency between studies' results, which is calculated based on the Q statistic [100%x(Q-df)/Q; Higgins et al., 2003). It scales from 0-100%, with low, moderate and high values of 25%, 50% and 75% respectively (Higgins et al., 2003).

Moderator Analysis

Studies included in the meta-analysis were coded for four potential categorical moderators (each discussed next). Moderator analysis was conducted in *metafor* using the '*mods*' function with each sub-category of moderator. This function conducts meta-regression analysis and allows the user to test whether the proposed categorical moderator produces a statistically significant effect on the outcome.

GBG. The first moderator specified whether or not the game applied was described in the method by the authors as a version of the GBG. This included positive iterations of the GBG and GBG modifications. This moderator was included based on the enduring popularity of the GBG and its versions compared to other game-based interventions. It was expected that a large proportion of the games identified would be versions of the GBG (based on the number of studies identified in previous reviews e.g., Bowman-Perrott et al., 2016). It was therefore important to consider the relative effectiveness of the GBG compared to other game-based interventions.

Grade Level. The second potential moderator was the grade level of the participants taking part in the study. Grade level was coded as primary (kindergarten-5th grade), secondary (6th-12th grade) or both. This categorisation was based on the tradition that students in the US traditionally begin middle school in 6th grade, and other similar reviews following a similar convention (Bowman-Perrott et al., 2016). Although Bowman-Perrott et al. (2016) did not find a significant difference for effects of the GBG between secondary and primary grades, this review would consider interventions other than the GBG, and therefore the potential for differing effects across grade levels was considered.

Sample Description. The third potential moderator was evaluated at two levels of sample description: individual or group. This referred to whether data was collected on the outcome as a group aggregate (i.e., a whole class group or sub-set of students within the class group) or as an individual outcome (i.e., one target student). This moderator was included based on the concern that sometimes an intervention may appear effective at the group level but does not impact behaviour as strongly at the individual level. This has important implications for teacher-reported social validity of an intervention, such that if and intervention is generally effective but one student is a non-responder and tends to disrupt class regularly, the teacher may not appreciate the general effects of the intervention.

Positive Reinforcement. The final moderator was whether the game-based intervention included positive reinforcement throughout the game in the form of provision of points. This was coded at two levels: positive reinforcement present versus not present. Some interventions do not include the use of positive reinforcement in the form of points throughout the game, but do reward students for engaging in low rates of target behaviours (i.e., DRL). However, for the purposes of examining this moderator, positive reinforcement was only considered present if points were administered rather than fouls/marks. For example, the traditional GBG (Barrish et al., 1969) involves the provision of marks/fouls rather than points so would be coded as ‘not present’.

Results

Study Design Quality Appraisal

Forty-nine studies, including 395 experiments, were evaluated using the WWC design standards (WWC, 2017) to assess the methodological quality of the study. Upon evaluation, 14 studies met the design standards and 16 studies met the design standards with reservations. Within these studies, 55 experiments met the design standards fully, 114 experiments met the standards with reservations and 226 experiments did not meet the standards. The remaining 19 studies did not meet WWC design standards. An outline of each study’s rating during the design evaluation phase and the specific standard or standards not met are outlined in Table 2.2.

Table 2.2.

Study Design Evaluation

Author (Year)	No. Experiments	DS#1	DS#2	DS#3	DS#4	Confound	No. Experiments Meeting Standards	No. Experiments Meeting Standard with Reservations	No. Experiments not Meeting Standards	Overall Rating
Barrish et al., (1969)	2	Y	N	Y	Y	None	0	0	2	NMS
Beeks & Graves (2016)	3	Y	N	Y	Y	None	0	0	3	NMS
Bostow & Geiger (1976)	1	Y	N	Y	Y	None	0	0	1	NMS
Casados (2012)	2	Y	Y	Y	Y	None	2	0	0	MS
Collier-Meek et al., (2017)	2	Y	Y	N	Y	None	0	0	2	NMS
Conklin et al., (2017)	60	Y	Y	Y	Y	None	0	50	10	MSR

Author (Year)	**No. Experiments**	**DS#1**	**DS#2**	**DS#3**	**DS#4**	**Confound**	**No. Experiments Meeting Standards**	**No. Experiments Meeting Standard with Reservations**	**No. Experiments not Meeting Standards**	**Overall Rating**
Dadakhodjaeva (2017)	3	Y	Y	N	Y	None	0	0	3	NMS
Dadakhodjaeva et al., (2019)	4	Y	Y	Y	Y	None	4	0	0	MS
Darch & Thorpe (1977)	1	Y	N	Y	Y	None	0	0	0	NMS
Davies & Witte (2000)	8	Y	N	Y	Y	Data collection methods inconsistent across baseline and intervention phases	0	0	8	NMS
Dillenburger & Coyle (2019)	105	Y	N	Y	Y	None	0	0	105	NMS
Donaldson et al., (2017)	12	Y	Y	Y	Y	None	10	1	1	MS
Donaldson et al., (2018)	4	Y	Y	Y	Y	None	0	4	0	MSR

Author (Year)	**No. Experiments**	**DS#1**	**DS#2**	**DS#3**	**DS#4**	**Confound**	**No. Experiments Meeting Standards**	**No. Experiments Meeting Standard with Reservations**	**No. Experiments not Meeting Standards**	**Overall Rating**
Donaldson et al., (2015)	10	Y	Y	Y	Y	None	0	6	4	MSR
Fallon et al., (2018)	2	Y	Y	Y	Y	None	0	2	0	MSR
Floress et al., (2017)	1	Y	N	Y	Y	None	0	0	1	NMS
Ford (2017)	8	Y	Y	Y	Y	None	8	0	0	MS
Grandy et al., (1973)	2	Y	N	Y	Y	None	0	0	2	NMS
Hansen et al., (2017)	1	Y	Y	Y	Y	None	1	0	0	MS
Hartman & Gresham (2016)	3	Y	Y	Y	Y	None	1	0	2	MS
Hernan et al., (2019)	6	Y	Y	Y	Y	None	6	0	0	MS
Hine et al., (2015)	8	Y	Y	Y	Y	None	0	8	0	MSR

Author (Year)	**No. Experiments**	**DS#1**	**DS#2**	**DS#3**	**DS#4**	**Confound**	**No. Experiments Meeting Standards**	**No. Experiments Meeting Standard with Reservations**	**No. Experiments not Meeting Standards**	**Overall Rating**
Hirsch et al., (2016)	1	Y	Y	Y	Y	None	0	1	0	MSR
Hoff & Ervin (2013)	2	Y	Y	Y	Y	None	2	0	0	MS
Kamps et al., (2015)	10	Y	N	Y	Y	None	0	0	10	NMS
Kamps et al., (2011)	6	Y	N	Y	Y	None	0	0	6	NMS
Kelshaw-Levering et al., (2000)	1	Y	Y	Y	Y	None	0	1	0	MSR
Kleinman & Saigh (2011)	3	Y	Y	Y	Y	None	0	3	0	MSR
Lynne et al., (2017)	6	Y	Y	Y	Y	None	6	0	0	MS
McGoey et al., (2010)	3	Y	N	Y	Y	None	0	0	3	NMS
McKissick et al., (2010)	2	Y	N	Y	Y	None	0	0	2	NMS

Author (Year)	No. Experiments	DS#1	DS#2	DS#3	DS#4	Confound	No. Experiments Meeting Standards	No. Experiments Meeting Standard with Reservations	No. Experiments not Meeting Standards	Overall Rating
Medland & Stachnik (1972)	10	Y	N	Y	Y	None	0	0	10	NMS
Mitchell (2014)	15	Y	Y	Y	Y	None	0	15	0	MSR
Mitchell et al., (2015)	3	Y	Y	Y	Y	None	0	2	1	MSR
Mitchem et al. (2001)	4	Y	Y	Y	Y	None	4	0	0	MS
Mudgal (2006)	4	Y	Y	N	Y	None	0	0	4	NMS
Naylor et al., (2018)	7	Y	Y	Y	Y	None	0	1	6	MSR
Nolan et al., (2014)	3	Y	Y	Y	Y	None	0	3	0	MSR
Patrick et al., (1998)	2	Y	Y	Y	Y	None	0	2	0	MSR
Radley et al., (2016)	9	Y	Y	Y	Y	None	3	6	0	MS

Author (Year)	No. Experiments	DS#1	DS#2	DS#3	DS#4	Confound	No. Experiments Meeting Standards	No. Experiments Meeting Standard with Reservations	No. Experiments not Meeting Standards	Overall Rating
Robichaux & Gresham (2014)	1	Y	Y	Y	Y	None	0	1	0	MSR
Saigh & Umar (1983)	3	Y	N	Y	Y	None	0	0	3	NMS
Speight (2018)	1[a]	Y	Y	Y	Y	None	1	0	0	MS
Tanol et al., (2010)	8	Y	Y	Y	Y	None	3	3	2	MS
Thorne & Kamps (2008)	24	Y	N	Y	Y	None	0	0	24	NMS
Vidoni et al., (2014)	1	Y	Y	Y	Y	None	0	1	0	MSR
Wahl et al., (2016)	8	Y	N	N	Y	None	0	0	8	NMS
Wills (2002)	6	Y	Y	Y	Y	None	0	4	2	MSR
Wright & McCurdy (2012)	4	Y	Y	Y	Y	None	2	2	0	MS

Note. No.= Number; Y= Yes Standard is Met; N= No Standard is Not Met; MS= Meets Standards; MSR= Meets Standards with Reservations; NMS= Does Not Meet Standards. [a]Note that there are additional experiments on individual students here, however because there are uneven numbers in the MBDs, we decided to omit. It is unclear which students should be compared with each other in the MBD.

Study Characteristics

Data were extracted only from studies meeting the design standards with or without reservations (*n* = 30). It should be noted that, where possible, data extracted only came from experiments within the studies that met the design standards with or without reservations (*n*= 169). Experiments referred to one participant and one outcome for a reversal/withdrawal or alternating treatments design or one outcome across a number of participants in a multiple baseline design. An example of this occurred in the study by Donaldson et al., (2017); 12 individual students took part in the study where the GBG was applied using a withdrawal design. Data were collected on disruptive behaviour across all 12 individuals, resulting in 12 experiments. One experiment did not meet the WWC design standards (WWC, 2017), however the remaining 11 did meet the standards. Therefore, when recording sample size, it was recorded as 11 rather than 12. Furthermore, in studies such as the Donaldson et al., (2017) study, where individuals or sub-groups of target children were monitored rather than the whole group, the sample size refers to the target children only (i.e., relating to the data which appears on the graphs). Donaldson et al., (2017) applied the GBG with 69 students across 3 classes, however data were collected on individuals within these classes rather than for each child in the class. Therefore, the sample size is reported as 11 rather than 69.

Student/School Characteristics

Of the 30 studies meeting the WWC standards with or without reservations (WWC, 2017), all studies except one were conducted in the USA (96.67% of studies). The study conducted outside of the USA was conducted in Belize (Nolan et al., 2014). Twenty-five studies were journal articles and the remaining five were doctoral theses. The total number of students taking part across all the studies was 1439. Of these students, 582 were reported as being female, 598 were reported as being male however six studies did not report the total number of males/females across the group of students (n = 259 not reported). One study did not report exact sample size (Speight, 2018). In some studies, where group and individual data were collected, the authors reported the whole sample size, but only specified the gender of the target students. For example, Conklin et al. (2017) specified the whole sample size as 80, and stated that of 13 targets, two were female, without specifying the number of females within the classes. Where this occurred, gender make-up was coded as ‘not reported’. Thirteen studies reported a larger proportion of males than females, nine reported a larger proportion of females than males and one reported an equal number. A range of ethnicities were reported across samples however, 16 studies did not report the sample's ethnicity. One further study only reported the ethnicity of target students (Conklin et al., 2017), who were all Hispanic. In cases where ethnicity for the entire sample was reported, (n = 722 students), 43.1% identified as African American, 21.1% identified as Hispanic and 5.7% identified as White. The remaining 30.2% were biracial, Latino, Mestizo, Kriol, Spanish, North American, Mayan, Native American or 'Other'. Many studies who did not report the ethnicity of the study sample reported the ethnicity of the students across the whole school, however this data were not coded in the current review. Ethnicity data should therefore be interpreted with extreme caution.

Three studies reported individual data only, 22 reported on group data only and five reported both group and individual data. Class groups taking part ranged from Kindergarten to 12th grade. Nineteen studies reported on primary/elementary grades only (Kindergarten to fifth grade), eight reported on secondary grades only (i.e., sixth grade plus) and three studies reported on a combination of both. The median grade level across studies was fourth grade (approximately equivalent to fourth class in an Irish setting). Further information on student characteristics by study are provided in Table 2.3.

Study Methods and Intervention Implementation

When the studies included in the review were examined to see what game-based intervention was implemented, it was found that ten studies evaluated the GBG only, in a manner close to the original format (i.e., students received marks for rule breaking; Barrish et al., 1969) and a further three studies evaluated a 'positive version' of the GBG (i.e., points awarded for positive behaviour rather than marks for disruptive behaviour). A further two studies evaluated the GBG and CBGG together (one was a comparison study and one study allowed teachers to choose between the two games). Two further studies evaluated other modifications of the GBG (e.g., the GBG with response cost, the GBG with a dependent group contingency). In total, therefore, 17 of the 30 included studies evaluated some version of the GBG with mainstream school students. Five studies evaluated the Class-wide Function-related Intervention Teams (CW-FIT) intervention. This intervention is similar to positive iterations of the GBG (i.e., the CBGG) but also involves a positive skills teaching component and points are awarded for use of these skills. The remaining eight studies involved implementation of the following games: Keep Busy and Carry On, a Self-management Game, a Group Contingency game, the Classwide Peer-assisted Self-management programme (CWPASM), the Quiet Classroom game, the Mystery Motivator game, the Fair Play game and the Praise game. Twenty-six of the studies involved the application of an interdependent group contingency and awarded points and/or marks to teams for following or breaking rules. A further two studies evaluated both interdependent and dependent group contingencies as part of a game-based intervention (Hartman & Gresham, 2016; Kelshaw-Levering et al., 2000). One study evaluated a dependent group contingency only (Vidoni et al., 2014) and one study did not incorporate a group contingency (Wills, 2002).

Dependent variables evaluated across studies were forms of academic engagement/on-task behaviour (i.e., behaviour targeted for increase) or disruptive behaviour (i.e., behaviour targeted for decrease). Twelve studies evaluated the effects of game-based interventions on disruptive behaviour only, six studies evaluated academically engaged/on-task behaviour only and twelve studies evaluated both. Other behaviours examined which are linked to the management of classrooms, included mobile device presence (Hernan et al., 2019), latency to on-task behaviour (Hine et al., 2015) and decibel level (Radley et al., 2016) which all could be considered to be disruptive classroom behaviours, in that they were targeted for reduction. Dependent variables for each study are outlined in Table 2.3.

In 26 studies, the class teacher was the sole intervention agent. In one study, an intervention specialist was available in the classroom to implement the intervention independently of the teacher (Hernan et al., 2019). In one other study, the researcher/experimenter was the sole intervention agent (Donaldson et al., 2015). One study evaluated the differential effects of different implementers across different phases, therefore the experimenter, teacher and individual students implemented the game-based intervention at different times throughout the study (Donaldson et al., 2018). One other study involved the experimenter and teacher implementing the game at different times (Donaldson et al., 2017). All but one study reported on rewards used and, in the study where it was unclear (Mitchem et al., 2001), it is highly likely rewards were used given the teams and points components. The range of rewards used varied from study to study, but were generally age appropriate tangibles (e.g., stickers, stamps, small edibles, a choice from a box of small tangible prizes) or non-tangibles (e.g., free time, extra recess). Twenty-six studies reported collecting some form of treatment integrity data during at least some data collection sessions. Twenty-four studies reported on teacher social validity, almost all of which reported positive teacher responses and no study reported completely negative teacher responses. Seventeen studies reported on student social validity, all of which reported positive outcomes. Key characteristics of the 30 included studies, including a brief description of the game-based intervention under investigation, are outlined in Tables 2.3. and 2.4. Table 2.3. includes key characteristics relating to the study samples, setting and dependent variables and Table 2.4. includes key characteristics relating to the independent variables, that is, the game-based intervention evaluated.

Table 2.3.

Key Characteristics of Studies Meeting the WWC Standards With and Without Reservations

Author (Year)	Exp. MS/MSR	Design	Sample Size	Sample Ethnicity	Grade Level	Subject	DV (Measurement)
Casados (2012)	2	MBD Across Participants	135	11.11% Anglo, 74.07% Hispanic, 11.85% Asian, 2.22% African American, .74% Indian	5th grade, 6th grade, 7th grade	Maths	Off-task behaviour (partial interval recording)
Conklin et al. (2017)	50	Reversal/withdrawal (various iterations depending on experiment, mostly ABAB)	80	n/r for Group, All target students Hispanic	KG, 2nd grade, 7th grade	KG= centres or reading; 2nd grade= Math; 7th grade= Science	On-task behaviour (momentary time sampling), compliance (percentage of opportunities), out-of-seat behaviour, talking out, hand-raising (partial interval recording)
Dadakhodjaeva et al. (2019)	4	MBD Across Participants	59	Whole Classes 98.3% African American, 1.7% Biracial (African American/Caucasian) Target Students 100% African American	KG	Small group activities with varying contexts across groups	Disruptive behaviour, academic engagement (momentary time sampling)

Author (Year)	Exp. MS/MSR	Design	Sample Size	Sample Ethnicity	Grade Level	Subject	DV (Measurement)
Donaldson et al. (2017)	11	Reversal/withdrawal (ABAB)	11	n/r	KG, 1st grade	Whole group instruction (sitting on a carpet, listening to the teacher, responding to questions and asking questions) Typically involved letter recognition (KG), reading and reading comprehension, story sequencing and math.	Disruptive behaviour (frequency/responses per min)
Donaldson et al. (2018)	4	Reversal/withdrawal (different combinations of A-B1-A-B2-A-B3-A-Choice; B phases varied in the implementer only- see IV description)	58	n/r	KG, 1st grade	Whole Group instruction (morning)	Disruptions (frequency/responses per min)
Donaldson et al. (2015)	6	Reversal/withdrawal (ABABAB for one experiment and ABAB for all others)	81	n/r	KG	Whole Group Instruction: Maths or literacy and instructions for small-group centers	Disruptions (frequency/responses per min)

Author (Year)	Exp. MS/MSR	Design	Sample Size	Sample Ethnicity	Grade Level	Subject	DV (Measurement)
Fallon et al. (2018)	2	MBD Across Participants	78	7.69% White, 44.87% Black/African American, 33.33% Latino, 6.41% Other	4th grade, 5th grade	Mathematics	Academic engagement, disruptive behaviour (momentary time sampling)
Ford (2017)	8	Reversal/withdrawal (ABAB in 3 classrooms, ABACAB in 1 classroom)	66	95.5% African American, 4.5% Hispanic	7th grade, 8th grade	English	Academically engaged behaviour, disruptive behaviour (momentary time sampling)
Hansen et al. (2017)	1	MBD Across Participants	74	n/r	2nd, 3rd & 4th grade	n/r	On-task behaviour (momentary time sampling)

Author (Year)	**Exp. MS/MSR**	**Design**	**Sample Size**	**Sample Ethnicity**	**Grade Level**	**Subject**	**DV (Measurement)**
Hartman & Gresham (2016)	1	Alternating Treatments Design	17	70.59% White (non-Hispanic), 23.53% African American, 5.88% Hispanic	KG	Morning Meeting and Calendar math time	Rule violations (frequency/responses per min)
Hernan et al. (2019)	6	Alternating Treatments Design	43	79% African American, 14% Caucasian, 7% Hispanic	9th Grade	Algebra	Academic engagement (momentary time sampling), off-task behaviour, mobile device presence (partial interval recording)
Hine et al. (2015)	8	Reversal/withdrawal (ABAB & ABAB plus Fading)	4	n/r	1st grade, 2nd grade, 3rd grade	Centers-based activities	Latency to on-task behaviour (latency), on-task behaviour (momentary time sampling)
Hirsch et al. (2016)	1	Reversal/withdrawal (ABAB)	20	n/r	2nd grade	PE	Small-group engagement (momentary time sampling)
Hoff & Ervin (2013)	2	MBD Across Participants	64	N/r	2nd grade	Maths & Reading	Disruptive behaviour (partial interval recording)
Kelshaw-Levering et al. (2000)	1	Reversal/withdrawal (ABACBC)	4	n/r	2nd grade	Random across day	Disruptive behaviour (partial interval recording)

Author (Year)	Exp. MS/MSR	Design	Sample Size	Sample Ethnicity	Grade Level	Subject	DV (Measurement)
Kleinman & Saigh (2011)	3	Reversal/withdrawal (ABAB+ Follow-up)	26	23% African American, 73% Hispanic, 4% did not report ethnicity	9th grade	History	Talk/verbal disruption, aggression/physical disruption, seat-leaving (partial interval recording)
Lynne et al. (2017)	6	Reversal/withdrawal (ABAB)	65	n/r	1st grade, 4th grade	n/r	Academically engaged behaviour, disruptive behaviour (momentary time sampling)
Mitchell (2014)	15	Reversal/withdrawal (ABAB)	63	85.7% African American, 11.1% Biracial (African American/Caucasian), 1.6% Caucasian, 1.6% Hispanic	10th, 11th & 12th grade	World History, English, US History	Academic engagement (momentary time sampling), disruptive behaviour, inappropriate vocalisations, out of seat, inappropriate touching (partial interval recording)

Author (Year)	**Exp. MS/MSR**	**Design**	**Sample Size**	**Sample Ethnicity**	**Grade Level**	**Subject**	**DV (Measurement)**
Mitchell et al. (2015)	2	Reversal/withdrawal (ABAB)	44	90.9% African American, 4.55% Biracial (African American/Caucasian), 2.27% Hispanic, 2.27% Caucasian	9th, 10th, 11th & 12th grade	Transitions to Algebra, Spanish 2	Disruptive behaviour (partial interval recording)
Mitchem et al. (2001)	4	MBD Across Participants	97	Whole Classes 22.68% Hispanic, Target Students 77.78% Caucasian, 22.22% Hispanic	7th grade	Language arts	On-task behaviour (duration for whole group; whole interval recording for target students)
Naylor et al. (2018)	1	Reversal/withdrawal (ABAB)	11	n/r	1st grade	Science and social studies focus (Core Knowledge Curriculum)	On-task behaviour (momentary time sampling)

Author (Year)	Exp. MS/MSR	Design	Sample Size	Sample Ethnicity	Grade Level	Subject	DV (Measurement)
Nolan et al. (2014)	3	Reversal/withdrawal (ABAB with follow up - there was a multiple baseline component but for purposes of the review, it is considered as a reversal design)	32	19% Mestizo, 10% Kriol, 29% Spanish, 32% North American, and 10% Mayan	Beginners (i.e. KG), Infant II (i.e. 2nd grade), and a combined Standard II and III classroom (i.e. 4th and 5th grade)	Maths (Classrooms 1 & 3), Language Arts (2)	Problem behaviour (partial interval recording)
Patrick et al. (1998)	2	MBD Across Participants	67	n/r	4th grade, 5th grade, 6th grade	Physical Education (volleyball)	Appropriate social behaviours, Inappropriate social behaviours (frequency)

Author (Year)	Exp. MS/MSR	Design	Sample Size	Sample Ethnicity	Grade Level	Subject	DV (Measurement)
Radley et al. (2016)	9	Reversal/withdrawal (ABAB- there was a multiple baseline component but for purposes of the review, it is considered as a reversal design)	56	96.5% African American, 3.6% Hispanic	1st grade	Maths, Student-directed language arts centers, group instruction	Academically engaged behaviour, disruptive behaviour (momentary time sampling), noise level (Decibel 10^{th} iPad application)
Robichaux & Gresham (2014)	1	MBD Across Participants	67	n/r	1st grade, 2nd grade, 3rd grade	Ms Swan: Afternoon math class; Ms Black: Morning work period; Ms Cullen: Morning reading lesson	Talking without permission (all three classrooms); Out of Seat behaviour (Ms Black only) (frequency per min)
Speight (2018)	1	MBD Across Participants	n/r	n/r	6th grade, 7th grade	Language Arts, Science	On-task behaviour (momentary time sampling)
Tanol et al. (2010)	6	Reversal/withdrawal (ABACBC & ACABCB)	6	100% Native American	KG	Morning Meeting	Rule violations (partial interval recording)
Vidoni et al. (2014)	1	MBD Across Participants	70	n/r	6th grade	Physical Education	Steps per min (i.e., engagement with PE class; pedometer)

Author (Year)	**Exp. MS/MSR**	**Design**	**Sample Size**	**Sample Ethnicity**	**Grade Level**	**Subject**	**DV (Measurement)**
Wills (2002)	4	MBD across participants & reversal/withdrawal (ABAB)	4	n/r	2nd grade, 3rd grade	Reading instruction, maths	On-task behaviour (momentary time sampling), inappropriate behaviour (frequency; mean per 10 min observation)
Wright & McCurdy (2012)	4	Reversal/withdrawal (ABAC & ACAB)	37	n/r	KG, 4th grade	Language arts class-independent seat work	On-task behaviour, disruptive behaviour (combination of momentary time sampling & partial interval recording)

Note. Exp. MS/MSR= Experiments Meeting Standards/Meeting Standards with Reservations; MBD = Multiple Baseline Design, KG = Kindergarten, DV = Dependent variable

Table 2.4.

Key Characteristics of the Independent Variables in Studies Meeting the WWC Standards With and Without Reservations

Author (year)	IV/Game Evaluated	IV Description
Casados (2012)	GBG	Teams, rules reviewed; mark for team when team member breaks a rule; periodic praise for good behaviour; winning teams identified on the board (those remaining below pre-set criterion) and rewards distributed
Conklin et al. (2017)	CW-FIT	Three components: 1. Taught students appropriate replacement behaviours for current problem behaviours plus additional skills. 2. Taught students to ignore other's inappropriate behaviour. 3. Differential reinforcement including frequent attention for appropriate behaviour, incorporating a group contingency token economy whereby teams worked towards a point criterion (similar to CBGG)
Dadakhodjaeva et al. (2019)	GBG	Teams; rules reviewed; mark for team when team member breaks a rule; marks reviewed at end of game and teams not exceeding criterion received a prize
Donaldson et al. (2017)	GBG	Teams; rules reviewed; mark for team when a team member breaks a rule; teams scoring less marks than criterion eligible for reward. If all teams exceeded the criterion; then the team with the least amount of points would win.
Donaldson et al. (2018)	GBG	Teams; rules reviewed; game led by teacher, experimenter or student depending on phase; mark given to teams when a member broke a rule; teams remaining below criterion eligible for a prize. If all teams exceeded the criterion, the team with the lowest amount of marks received a prize.
Donaldson et al. (2015)	GBG	Teams; rules reviewed; teams received a mark when a member broke a rule; teams remaining below the criterion received a prize immediately after the game ended.
Fallon et al. (2018)	GBG & CBGG	Teams; rules reviewed; marks for teams where members displayed inappropriate behaviour (GBG) or points for appropriate behaviour (CBGG); reason for mark/point provided by teacher; team with the least amount (GBG) or most amount (CBGG) of marks/points won the game and verbal praise provided to winners.
Ford (2017)	GBG Positive Version	Teams; rules reviewed; teams received points if following rules when teacher scanned the room (VI 2 min schedule). Points were awarded via Class Dojo. The team with most points eligible for a prize, or both teams if both met criterion. Two rewards drawn and teams with most points voted on reward.

Author (year)	IV/Game Evaluated	IV Description
Hansen et al. (2017)	CW-FIT	Skills displayed, precorrection at beginning of session. Team point chart displayed, and a daily point goal posted. Students given self-management charts and prompted to award points to selves if on task every 2-5 mins. Points awarded to teams for use of skills. Points tallied at the end of class and those reaching a criterion are given a reward.
Hartman & Gresham (2016)	GBG (interdependent vs dependent group contingency)	1. Interdependent version: Teams, rules reviewed, marks given to teams when a member broke a rule, mystery prize for teams remaining below criterion. 2. Dependent version: Same as interdependent version but team captains chosen for individual monitoring. Students not informed who these captains were until end of game. Teams whose captains remained below the criterion received a mystery prize.
Hernan et al. (2019)	GBG Positive Version	Teams; rules posted on a poster; teams received 15 bonus points if the whole team placed their phones in the 'Clear Box' and left them there. Teams got points every 10-15 mins if they were on task. Points totalled at end of game. Team totals were then announced, and mystery criterion drawn from jar. Teams meeting this criterion were rewarded.
Hine et al. (2015)	Keep Busy and Carry On (KBCO)	Game automated by a video. The video would cue students to move from one centre to another and upon cueing, a transition timer was displayed letting students know how long the transition was taking. The transition criterion was randomised, and upon timer expiry, the teacher gave feedback and recorded points for teams who had met the criterion. When the transition was over, the video would proceed and cue students to check whether they were on task during the activity. At this point, the teacher would scan the room and award points to groups where all students were on task. Team points were coloured on to a visual display at the end of the day and every 2-3 extra recess could be earned as a prize.
Hirsch et al. (2016)	CW-FIT	Class divided into 4 teams (teachers' choice). Teams were told of CW-FIT expectations, told how to win points and the daily criterion. The teacher scanned the class on a variable interval 2-min schedule, awarding points to teams following expectations. Points were recorded on a bulletin board and paired with behaviour specific praise. Team totals were added up at the end of the game and teams reaching the criterion were awarded a prize.
Hoff & Ervin (2013)	Self-management Game	Teacher-directed phase: Rules reviewed, teacher rated class behaviour 1-5 at end of class and this corresponded with points awarded. Points

Author (year)	IV/Game Evaluated	IV Description
		could be swapped for a back-up reinforcer. Self-management Phase: Same as teacher-directed phase but students self-rated their own and the class behaviour. Teacher held a vote at end of class to determine class rating. A rule was randomly selected, and students reported their rating for that rule to the class. The majority rating became the class rating. The teacher then reported their own rating. If this matched the students rating the class got the amount of points plus a bonus point. If they were within one point of each other, the students kept their rating. If they were within two or more points of each other, no points were awarded.
Kelshaw-Levering et al. (2000)	Group Contingency Game	B phases (reward randomised): rules reviewed, group contingency applied across whole class, students given a 'check' upon engaging in problem behaviour. If class earned checks below a criterion, the whole class received a mystery reward. C phases (all components randomised): similar to B phases, however behaviour target was drawn randomly, the criterion was drawn randomly and whether the reward was based on individual or group behaviour was drawn randomly. If individual was drawn at this stage, then an individual's name was drawn randomly and finally, the prize was drawn randomly.
Kleinman & Saigh (2011)	GBG	Teams; rules reviewed; marks for teams on which a student broke a rule; team with fewest marks at the end of the game won a prize.
Lynne et al. (2017)	GBG Positive Version	Teams; rules reviewed; points awarded to teams following the rules using Class Dojo. Teams surpassing a criterion were eligible for a reward. A member of a winning teams randomly drew a reward slip from a jar.
Mitchell (2014)	GBG	Teams; rules reviewed; marks for teams on which a student broke a rule; team with fewest marks at the end of the game won a prize, or both teams won if neither reached the criterion.
Mitchell et al. (2015)	GBG	Teams; rules reviewed; marks for teams on which a student broke a rule; team with fewest marks at the end of the game won a prize, or both teams won if neither reached the criterion.
Mitchem et al. (2001)	Classwide Peer-assisted Self-management Programme (CWPASM)	Twelve steps reported: 1. Self-management defined and taught, 2. Definition of ABC's of SM provided plus examples, 3. Rules reviewed with role-play, 4. Students learned behaviour rating system and points procedure, 5. Students learn to evaluate their behaviour 6. Class rules, rating system and statements posted to the walls, 7. Students paired with a partner and pairs assigned to teams, 8. Teacher set a timer and students compared their perceptions of their behaviour with the rating system and recorded their response and rated their partner's behaviour

Author (year)	IV/Game Evaluated	IV Description
		when the tone sounded. 9. Partners compared their ratings and received points when ratings matched, 10. Points totalled after last tone sounded, 11. Teams surpassing previous day's points/team with the highest average were winners, 12. One random partnership selected by teacher and if their ratings matched the teacher's ratings, they won a 10-point bonus for their team.
Naylor et al. (2018)	CW-FIT	Applied across three components 1. Three appropriate behaviours/skills taught to students 2 & 3 Group contingency & Differential reinforcement: Students divided into teams and the teacher was prompted by a timer to award points to teams where students were displaying skills taught at step 1. Daily goal for points set at beginning of period and teams reaching this goal were eligible for a prize.
Nolan et al. (2014)	GBG	Teams; rules reviewed; marks given to teams on which a member broke a class rule. The team with the fewest marks at the end of the period won the prize. All teams could win if they all remained under a criterion. If both teams surpassed the criterion, neither team won.
Patrick et al. (1998)	GBG Modified Version	Teams (based on volleyball teams); points given to teams where members displayed appropriate social behaviours. Teams lost points for inappropriate social behaviours or 'false acts' of appropriate social behaviour. Students recorded their own points during scheduled breaks in the game. Teams meeting or exceeding the criterion (which was to exceed the previous day's criterion/ten times the baseline on the first day), received a reward. At the end of the unit, the team who met the criterion most often won an extra reward.
Radley et al. (2016)	The Quiet Classroom Game (QCG)	Rules reviewed at beginning of game. The students were told that the teacher would periodically check the noise levels during 15-min sessions via an iPad app. If the students as a group were at a certain noise criterion when the teacher carried out a check, the students were given feedback and a smiley face was placed in a box on the classroom whiteboard. If they didn't meet the criterion when the teacher carried out a check, the teacher gave relevant feedback and placed a frowning face in the box. If the class had at least 5 smiley faces by the end of the period, the class got a reward.
Robichaux & Gresham (2014)	Mystery Motivator Intervention	When a class rule was broken, the teacher recorded it on the board and provided feedback to the offending student, reminding them of a replacement behaviour. Marks were totalled at the end of the game and if the class met a criterion, the teacher coloured in the day on a chart. If an 'M' was revealed, the class was rewarded (M was written in invisible ink on some of the days). The mystery prize was written in an envelope

Author (year)	**IV/Game Evaluated**	**IV Description**
		and the class received that prize if an 'M' was revealed upon winning the game.
Speight (2018)	CW-FIT	Teams; teams earned points for following the rules; the teacher paired awarding of points with behaviour specific praise; teams meeting the criterion at the end of class had access to a prize.
Tanol et al. (2010)	GBG Modified Version	Response-cost versus reinforcement Response cost: teams; rules reviewed; teams started game with 4 stars and lost a star contingent on a team member breaking a rule. Teams with one star or more at the end were eligible for a prize. Reinforcement: teams; rules reviewed; teams started the game with no stars, and earned stars contingent on rule following. Teams with three or more start at the end were eligible for a prize. Weekly prizes for teams meeting criterion two or more days in a row.
Vidoni et al. (2014)	Fair Play Game	Teams; goal was to increase steps taken during PE. One unidentified student was chosen from each team at the end of the intervention period and if that student had reached the goal for number of steps, then the team won the prize.
Wills (2002)	The Praise Game	The participating student was given a page where they self-recorded incidences of teacher praise towards them during a certain period in the day. The student was given expectations for the duration of the game and the teacher praised them when they were following these expectations. When the student reached their goal for praise, they were eligible to receive a prize.
Wright & McCurdy (2012)	GBG & CBGG	Teams; start of game signalled. GBG: Mark given to teams where a member engaged in disruptive behaviour. Marks tallied at end of game and teams not exceeding a pre-set mystery criterion by the end of the game received a prize. CBGG: Points given to teams following rules when teacher scanned the room (VI 4 min cue). Teams surpassing a criterion of points eligible for a prize. Weekly mystery criterion in both versions where teams could earn a weekly prize.

Note. IV = Independent variable

Effect Size Calculation and Meta-analysis

As specified earlier in the Methods section, effect sizes were calculated based on experiments within studies meeting the WWC design standards with or without reservations (WWC, 2017). Effects were calculated for experiments individually and then they were combined for inclusion in a meta-analysis, such that the effects included in the meta-analysis are made up of various experiments within studies. A table outlining the exact experiments making up the various effect sizes included in the meta-analysis is included in the Appendix (Appendix C, Table C-1). This table provides clarity to the reader as to exactly what is contributing to the effect sizes presented here.

Tau

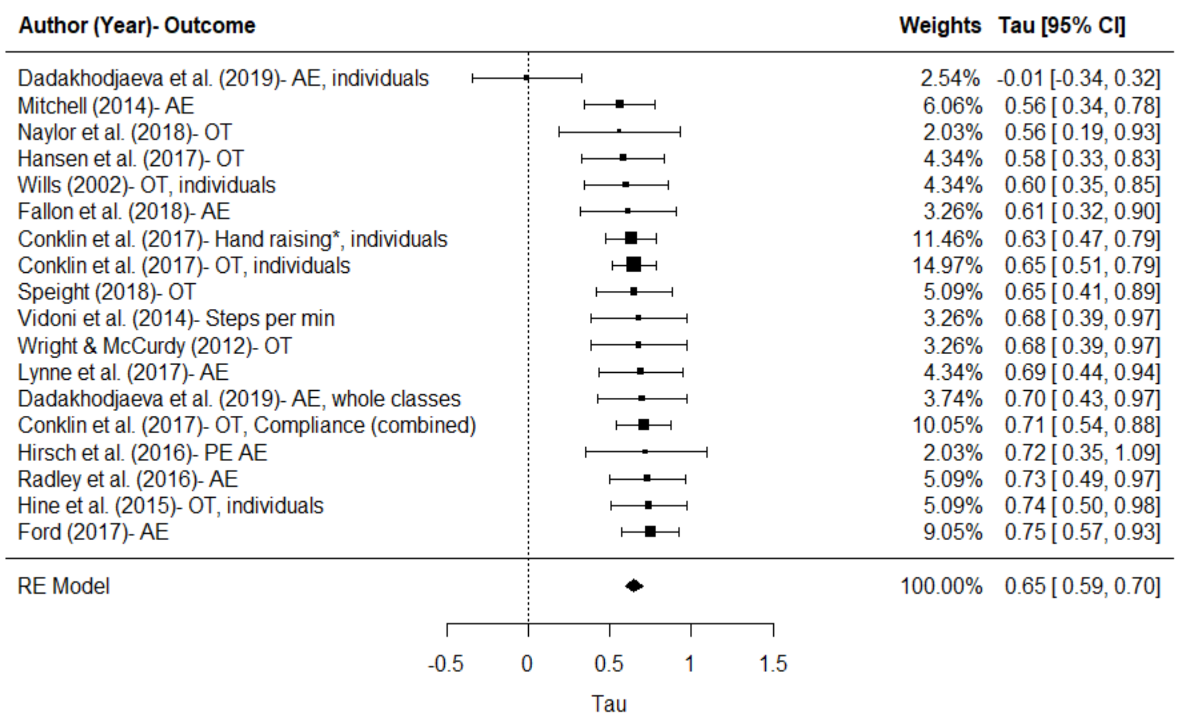
Tau effect sizes were calculated for 26 of the 30 included studies. Two studies were alternating treatment designs for which Tau is not appropriate (Hartman & Gresham, 2016; Hernan et al., 2019) and in two further studies (Mitchem et al., 2001; Patrick et al., 1998), it was not possible to extract the graphed data. Some studies evaluated the effects of a game-based intervention on positive behaviour only (e.g., academic engagement, on-task behaviour), some targeted disruptive behaviour only (e.g., rule violations) and some targeted both. There were more experiments where disruptive behaviours were targeted than positive behaviour, therefore, 18 effect sizes were calculated based on experiments aiming to increase the target behaviour (e.g., academic engagement) and 25 effect sizes were calculated based on experiments aiming to decrease the target behaviour (e.g., disruptive behaviour).

Positive Behaviours. Across 15 studies (resulting in 18 effect sizes), Tau effect sizes for positive behaviour ranged from small (-.01) to large (.75). Three of the 18 effects were between .56 and .58 (moderate) and 14 of the 18 effects were greater than .6 (large). One negative effect size was calculated and relates to an experiment where baseline correction was required during calculations (Dadakhodjaeva et al., 2019). One of the participant's behaviour was demonstrating an upward trend during baseline and when this was corrected, the effect size was negative.

The 18 Tau effect sizes for positive behaviour were input into the *metafor* package for R (Viechtbauer, 2010) and a random-effects meta-analysis was conducted to calculate an omnibus effect size. The omnibus effect size based on the random-effects model was .65 (SE = .03, p<.0001, 95% CI [.59, .70]), suggesting a large, significant positive effect of game-based classroom management interventions on positive student behaviour. The results indicated that there was no significant heterogeneity among the studies (Q [df = 17] = 19.70, p = .29) and there was low variance between the studies (I^2=0%). A forest plot outlining Tau and 95% CIs for individual studies and outcomes is presented in Figure 2.2.

Figure 2.2.

Forest Plot of 18 Tau Effect Sizes for Positive Behaviour Outcomes Across 15 Included Studies



Note: AE = Academic Engagement, OT= On-task Behaviour, PE= Physical Education.

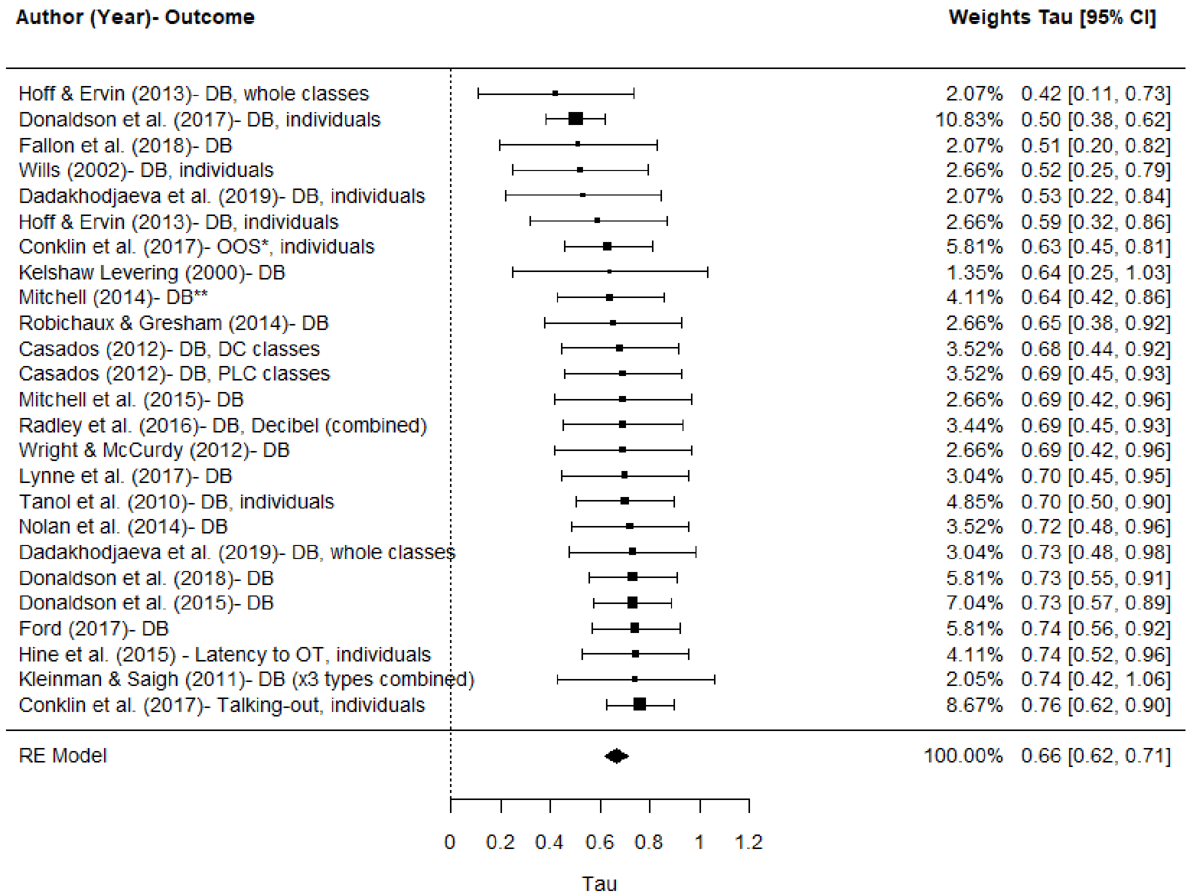
* Hand-raising and on-task behaviour were not combined as a synthetic variable as the same individuals do not make up the respective effect sizes for each (due to inability to extract data for some experiments).

Disruptive Behaviours. When considering Tau effect sizes for disruptive behaviour, all signs were changed to positive signs, such that a positive Tau value indicates a positive effect on behaviour (i.e., a decrease in behaviour). Across 21 studies (resulting in 25 effect sizes), Tau effect sizes for disruptive behaviour ranged from moderate (.42) to large (.76). Six of the 25 effects were between .42 and .59 (moderate) and the remaining 19 effects were greater than .6 (large).

The 25 Tau effect sizes for disruptive behaviours were input into the *metafor* package for R (Viechtbauer, 2010) and a random-effects meta-analysis was conducted to calculate an omnibus effect size. When conducting this analysis, the Fisher scoring algorithm was used as part of the REML estimator, and in this case it did not converge, resulting in an error message. The issue was rectified by decreasing the step length and did not reflect a deficiency in the estimator used (Viechtbauer, 2020). The omnibus effect size based on the random-effects model was .66 (SE = .02, $p < .0001$, 95% CI [.62, .71]), suggesting a large, significant positive effect of game-based classroom management interventions on disruptive student behaviour. The results indicated that there was no significant heterogeneity among the studies (Q [df = 24] = 18.32, $p = .79$) and there was low variance between the studies ($I^2 = 11.65\%$). A forest plot outlining Tau and 95% CIs for individual studies and outcomes is presented in Figure 2.3.

Figure 2.3.

Forest Plot of 25 Tau effect sizes for disruptive behaviour outcomes across 21 included studies



Note: DB = Disruptive Behaviour (all variables referring to disruptive behaviour generally are coded as DB here. For example, Tanol et al., record rule violations in their study which we code here as DB), OOS= Out-of-Seat, DC= Direct consultation, PLC= Professional Learning Community, OT= On-task**.** Data refers to group data unless otherwise specified.

*****OOS and talking-out were not combined as a synthetic variable as the same individuals do not make up the respective effect sizes for each (due to inability to extract data for some experiments). ******Only the composite DB variable was included in the meta-analysis as this is a combination of the 3 types graphed elsewhere in the study.

A reliable difference formula $[(L1-L2)/\ \mathrm{Sqrt}(SE1^2+SE2^2)]$, was used to assess whether or not there were differences between the impact of the game-based intervention on positive behaviours versus disruptive behaviour. This formula is based on a t-test and has been used in previous reviews to determine if different levels of moderators differed significantly from each other (Bowman-Perrott et al., 2016). A z-score is calculated using the above formula and a p-value can be calculated from this z-score. The effects of game-based interventions did not differ significantly between positive behaviour applications and disruptive behaviour applications (z = -.2774, p = .78).

BC-SMD

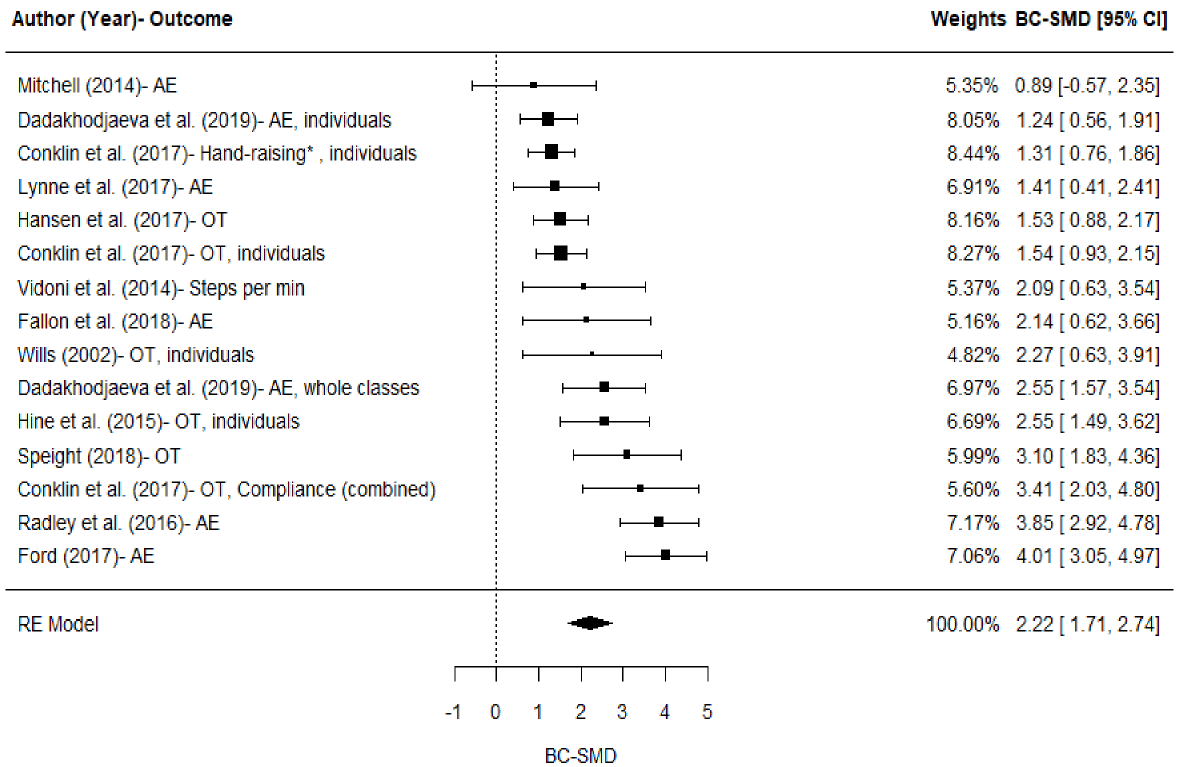
BC-SMD effect sizes were calculated for 20 of the 30 included studies. Studies for which BC-SMD effects were not calculated included those for which data could not be extracted (Mitchem et al., 2001; Patrick et al., 1998), alternating treatment designs (Hartman & Gresham, 2016; Hernan et al., 2019) and reversal/withdrawal designs conducted across less than three participants/participant groups (Hirsch et al., 2016; Kelshaw-Levering et al., 2000; Kleinman & Saigh, 2011; Mitchell et al., 2015; Naylor et al., 2018; Wright & McCurdy, 2012).

Positive Behaviours. Twelve studies contained experiments which allowed for calculation of BC-SMD effect sizes for positive classroom behaviours. Fifteen effects were calculated in total. These ranged from small (.89) to large (4.01). One effect size was small (.89), five were medium (1.24-1.54) and the remaining nine were large (2.09-4.01).

The effects and their variance were input into the *metafor* package for R (Viechtbauer, 2010) and a random-effects meta-analysis was conducted to arrive at an omnibus effect size. The omnibus effect size based on the random-effects model was 2.22 (SE = .26, p<.0001, *95% CI* [1.71, 2.74], indicating an overall significant, large effect for game-based interventions on positive classroom behaviours. There was however significant heterogeneity between the studies (Q[df = 14]= 60.14, p < .0001) and there was a substantial amount of variance detected (I^2= 77.38%). Figure 2.4. is a forest plot outlining the results of the random-effects meta-analysis.

Figure 2.4.

Forest plot of 15 BC-SMD Effect Sizes for Positive Behaviour Outcomes Across 12 Included Studies



Note: AE = Academic Engagement, OT= On-task Behaviour, PE= Physical Education. Data refers to group data unless otherwise specified.

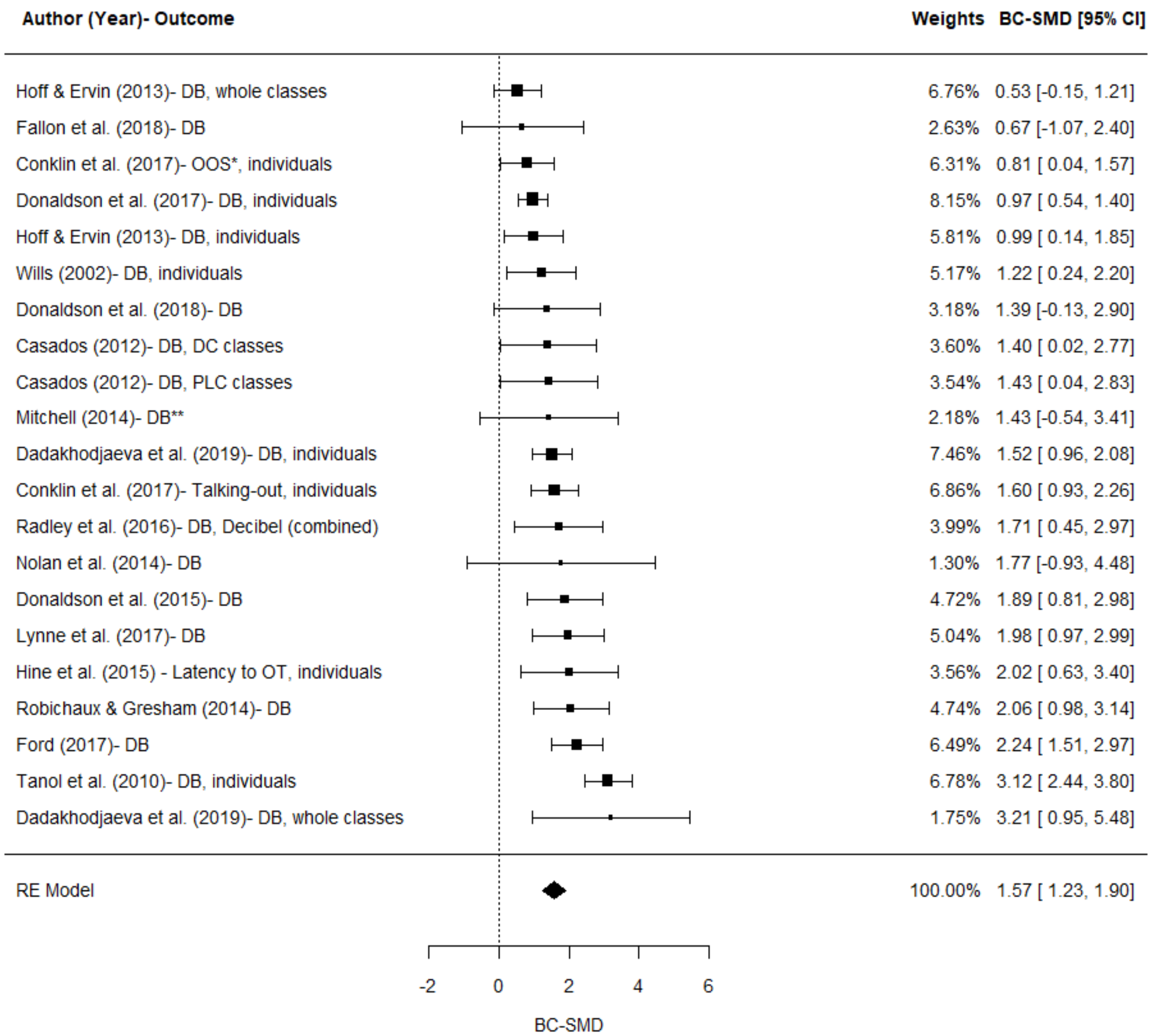
* Hand-raising and on-task behaviour were not combined as a synthetic variable as the same individuals do not make up the respective effect sizes for each (due to inability to extract data for some experiments).

Disruptive Behaviours. Seventeen studies contained experiments which allowed for calculation of BC-SMD effect sizes for disruptive behaviours. Twenty-one effects were calculated in total. These effect sizes ranged from small to large (range = .53- 3.21). Four of the effects were small (.53-.97), ten were medium (.99-1.77) and the remaining seven were large (1.89-3.21).

The effect sizes and their variance were input into the *metafor* package for R (Viechtbauer, 2010) and a random-effects meta-analysis was conducted to arrive at an omnibus effect size. The omnibus effect size based on the random-effects model was 1.57 (*SE* = .17, p<.0001, *95% CI* [1.23, 1.90], indicating an overall significant, medium effect for game-based interventions on disruptive classroom behaviours. There was however significant heterogeneity between the studies (Q [df = 20] = 50.39, p < .001) and there was a moderate amount of variance detected (I^2= 59.05%). Figure 2.5. outlines these results in a forest plot.

Figure 2.5.

Forest plot of 21 BC-SMD Effect Sizes for Disruptive Behaviour Outcomes Across 17 Included Studies



Note: DB = Disruptive Behaviour (all variables referring to disruptive behaviour generally are coded as DB here. For example, Tanol et al., record rule violations in their study which we code here as DB), OOS= Out-of-Seat, DC= Direct consultation condition, PLC= Professional Learning Community condition, OT= On-task**.**

Data refers to group data unless otherwise specified.

*****OOS and talking-out were not combined as a synthetic variable as the same individuals do not make up the respective effect sizes for each (due to inability to extract data for some experiments). ******Only the composite DB variable was included in the meta-analysis as this is a combination of the 3 types graphed elsewhere in the study.

The reliable mean difference formula was applied again here to assess whether there were statistical differences between the effects of game-based interventions on positive behaviour versus disruptive behaviours. The results indicated that there was a significant difference between the effects ($z = 2.09$, $p = .03$), such that larger effects were observed for positive behaviour outcomes versus disruptive behaviour outcomes.

Moderators

There was significant heterogeneity and a moderate to substantial amount of variance between studies included in the meta-analyses which considered the BC-SMD as the effect size. It was therefore appropriate to probe the potential effects of moderators. Four moderators were evaluated here: whether the game was specified as a version of the GBG or not, whether the outcome refers to the effect of a game-based intervention on primary grades (i.e., K-5), secondary grades (i.e., 6-12) or a combination of both, whether the sample being exposed to the intervention was an individual or group and whether positive reinforcement (the provision of points) was a key mechanism in the game. Across positive and disruptive classroom behaviours, none of the moderators yielded a statistically significant coefficient. These findings are summarized in Table 2.5.

Table 2.5.

Moderator Analysis Across Four Moderators Included in the Synthesis

Potential Moderator	Outcome Category	*QM*	*p*	I^2	R^2
GBG	Positive	.23	.63	78.43%	0%
	Disruptive	2.37	.12	54.34%	14.28%
Grade Level	Positive	.86	.65	77.24%	0%
	Disruptive	1.21	.55	58.62%	0%
Sample Description	Positive	2.67	.10	72.66%	18.68%
	Disruptive	.11	.74	60.08%	0%
Positive Reinforcement[a]	Positive	1.43	.23	78.68%	1.24%
	Disruptive	.18	.67	39.68%	0%

[a]Outcomes from Fallon et al., (2018) and Tanol et al., (2010) not incorporated in this analysis as they evaluated combinations of positive reinforcement, positive punishment and response cost.

Discussion

The purpose of the current review and meta-analysis was to systematically identify and synthesise single-case studies evaluating the effects of game-based interventions for classroom management related behaviours. Initially, 49 studies were identified which evaluated such interventions however, after appraisal of the study designs, 30 studies were included in the final synthesis. This is the first review to collate information on game-based classroom management interventions in one place and there are several important findings which will now be highlighted.

Game-based Interventions Present in the Literature

Data collected during study characteristic coding allowed for a distinct list and description of game-based interventions which have been applied in the literature to be developed. As evidenced in the Results section (see Table 2.4.), a majority of the 30 included studies evaluated the GBG as traditionally applied (*n* = 10), or some type of positive or modified GBG (*n* = 7). Of these studies, only one was carried out before 2010 (Patrick et al., 1998). This suggests that, despite novel game-based interventions making headway in the literature, the GBG and its variants have remained popular over the past decade. The enduring popularity of the GBG in research and practice is striking, considering it first appeared in the literature over 50 years ago. It is important to note that almost all of the high-quality evaluations of the GBG were conducted in the USA, with the exception of the study by Nolan et al., (2014; conducted in Belize). This cultural gap in the literature is surprising considering the GBG is a relatively simple intervention which has large positive effects on behaviour. There is a clear gap in our knowledge on the implementation of the GBG and its variants, and game-based interventions more generally in settings outside of the USA.

Chapter 1 outlined the discrepancies in language used to describe versions of the GBG, particularly when it is modified to involve a focus on positive behaviour (i.e., the awarding of points). These discrepancies were confirmed here, with different terminology used to describe positive versions of the GBG across studies. Of the seven studies evaluating positive or modified versions of the GBG, two referred to a 'positive version' as the CBGG (Fallon et al., 2018; Wright & McCurdy, 2012). Three further studies refer to a positive version of the GBG as the GBG but specify the positive modification in the study methods section (Ford, 2017; Hernan et al., 2019; Lynne et al., 2017). Where authors specified their positive iteration of the game as a modified GBG, one called it the GBG-reinforcement (Tanol et al., 2010) and one referred to it as a modified GBG (Patrick et al., 1998). Given the difference in principles applied in positive versions of the GBG versus the traditional GBG (outlined in Chapter 1), future research may consider adopting a name for positive versions of the GBG which distinguishes it clearly from the traditional GBG.

Following the GBG, the next most-applied game was the CW-FIT. The CW-FIT can be likened to the CBGG specifically given its focus on awarding points for appropriate behaviour in an interdependent group contingency format. The CW-FIT builds upon the CBGG by preceding the group contingency game with the explicit teaching of skills needed for engagement in prosocial and positive classroom behaviour. It is therefore a slightly more complex game-based intervention which may require more time and training for teachers. Similarly to the GBG, the CW-FIT has been applied solely in the USA, leaving gaps in current knowledge of its impact in diverse cultural settings.

Of the various other games evaluated (*n* = 8), all but one were based on a group contingency format with six incorporating an interdependent group contingency and one incorporating a dependent group contingency. It is evident that the GBG was an influence on the develop of some of these games. For example, Vidoni et al. (2014) refer to the Fair Play Game as a variation of the GBG in the introduction to their study. In the current review it was considered a game in its own right rather than a version of the GBG, as the GBG was not mentioned in the study method and the Fair Play Game used a different group contingency strategy to the GBG (a dependent group contingency). Radley et al. (2016) also make reference to the GBG in their discussions of the Quiet Classroom Game, noting its similarities and differences. This reiterates the influence the GBG has had on the state of the literature, even where the development of new games is concerned.

With the seminal study on the GBG (Barrish et al., 1969) incorporating one of the first examples of a group contingency in classroom management (see Maggin et al., 2012), it is interesting to note that of the 30 studies included in the review overall, 29 incorporated group contingencies of some kind. This finding, although perhaps not surprising, is an important one. It was expected that many of the studies included would comprise of group contingencies, however it was not expected that they would make up such a large majority. There is already a strong evidence base for the practice of group contingencies in classroom settings (Maggin et al., 2012; 2017). and the current review has identified an important sub-set of group contingency interventions (with the exception of one included study that did not incorporate a group contingency; Wills, 2002). It is likely therefore, that the group contingency characteristic is central to the games' success as behavioural classroom management interventions. Future research may look towards evaluating game-based group contingency interventions versus those not framed as a game and assert whether there are differences in effectiveness between the two.

Populations Studied

Game-based interventions have been applied with diverse populations, with a generally equal gender balance overall where gender was reported. One point of note is that game-based interventions do not appear to have been studied in mainstream single-sex settings. Although some studies reported the sample target students as being all male (e.g., Tanol et al., 2010), the gender make-up of the whole class is not reported. This indicates a gap in the literature for future study on the topic, particularly in cultures like Ireland where single-sex education is relatively common. Although various ethnicities took part in the studies, many studies did not report sample ethnicity which makes it difficult to draw conclusions on the spread of ethnicities taking part in the studies based on the data gathered during this review. The majority of students were African American where ethnicity was reported, but around half of the students' ethnicity is unknown. Applications of game-based interventions are spread across all grade levels, however there are more applications with primary school children (those in fifth grade or younger) compared with secondary school children (sixth grade or older). This is not surprising given that group contingency research and GBG research are more common with primary school children (Bowman-Perrott et al., 2016; Maggin et al., 2017). Of those studies conducted with older children, many have been published in the last five years (e.g., Ford, 2017; Hernan et al., 2019; Mitchell et al., 2015) and many maintain a focus on positive behaviour by awarding points for rule following rather than marks for rule breaking (e.g., Conklin et al., 2017; Ford, 2017; Hernan et al., 2019). This is a promising development which has the potential to align more closely with studies on primary school children in the coming years.

Behavioural Outcomes Targeted by Game-based Interventions

An aim of this review was to specify which behavioural outcomes had been targeted by game-based interventions. The synthesis uncovered a range of behaviours targeted by game-based interventions, however most were variations of disruptive behaviour and academic engagement, such as rule violations, talking-out and on-task behaviour. Some novel, more specific behaviour targets were identified such as mobile phone presence (Hernan et al., 2019), steps per min as an indicator of engagement during a PE class (Vidoni et al., 2014) and decibel/noise level (Radley et al., 2016). There is therefore clearly promise in the use of game-based interventions to target novel behaviours in the classroom that are alternatives to the more traditional academic engagement. Alternative indicators of disruption which can be assessed immediately and accurately with technology, such as decibel level, may also be useful when evaluating game-based interventions. In the study by Radley et al., (2016), decibel level measured by an iPad application was significantly and positively correlated with disruptive behaviour which was measured via momentary time sampling and presented as a percentage of intervals (r = .75; calculated in order to facilitate calculation of synthetic effect size during meta-analysis). Although this may not always be an appropriate measure of disruption (e.g., during a music lesson, during group work where talking is permitted), it may be useful where quiet independent work is underway.

Effectiveness of Game-based Interventions in Classroom Management

The current review also aimed to evaluate if game-based interventions are effective in targeting classroom management related behaviour. This aim was addressed through a series of random-effects meta-analyses of outcomes across categories of positive and disruptive behaviours. Tau effect sizes revealed mostly large effects for the impact of game-based interventions on positive and disruptive behaviours, and the omnibus effect sizes for both topographies of behaviour were large. No very large effect sizes were evident according to the threshold put forward by Vannest and Ninci (2015). However, the Tau effect sizes calculated here are more conservative than that specified by Vannest and Ninci (2015; i.e., Tau_{novlap} and Tau-U), as outlined earlier in the Methods section. These large effects indicate the promise of game-based interventions as classroom management tools. There was no significant difference between Tau effect sizes for positive behaviours versus disruptive behaviours, which contrasts with a recent finding by Bowman-Perrott et al., (2016). Bowman Perrott et al. reported that the GBG was more effective in targeting disruptive/off-task behaviours than attention to task/on-task behaviours. Bowman-Perrott et al., (2016) however, included studies in their meta-analysis that did not fully meet the WWC design standards. This meant that some studies of limited quality were included in the analysis, limiting conclusions which can be drawn from their effect size calculations. Bowman-Perrott et al. also used different conventions in calculating Tau and in collating study effects (i.e., a fixed-effects model). It is also important to note that this review, although including a high number of studies on the GBG, is not solely focused on the GBG and included various other game-based interventions.

The second pair of meta-analyses included BC-SMD effect size calculations for positive and disruptive behaviours. For positive behaviours, a majority of the effects calculated were large and for disruptive behaviours, a majority were spread across the medium to large categories. The omnibus effect was slightly smaller for disruptive behaviours (Omnibus BC-SMD = 1.57) compared to positive behaviours (Omnibus BC-SMD = 2.22) in contrast to the findings based on the Tau data. As well as this, there was a significant difference between the omnibus BC-SMD values for positive behaviour versus disruptive behaviour. It must be noted that the studies and outcomes included in the BC-SMD meta-analyses were not identical to those included in the Tau meta-analyses, due to limits on the type of studies for which BC-SMD could be calculated. When these results are compared to a recent study examining the impact of group contingency interventions on challenging classroom behaviour, they are strikingly similar. Maggin et al., (2017) specified a medium effect for group contingencies on disruptive behaviours (1.56) and a slightly larger effect on academic engagement (1.88) based on the gradual effects model (i.e., similar to the model applied and reported here). This is not surprising, given this review has established that when game-based interventions are applied in mainstream classrooms, they are almost always applied using a group contingency.

Potential Moderators of Game-based Intervention Effectiveness

It is evident that a majority of the high-quality studies included in this review have been applied with primary-aged students (*n* = 19). However, through moderator analysis on the BC-SMD data, it was apparent that game-based interventions were effective across studies recruiting primary, secondary and a combination of both grade levels. This is consistent with previous research which found no significant difference between primary and secondary grades exposed to the GBG (Bowman-Perrott et al., 2016), or to group contingencies (Maggin et al., 2017). It must be reiterated here however, that the moderator analysis was conducted on studies for which the BC-SMD effect size could be calculated only, therefore some studies evaluating the effects of game-based interventions on classroom behaviour were omitted.

Research on game-based interventions has looked at whether interventions are as effective at the individual level as at the group level (e.g., Dadakhodjaeva et al., 2019), and therefore sample description was evaluated as a moderator. Sample description did not moderate results in this analysis, similar to a recent finding in Maggin et al's., (2017) group contingency review. This finding should not preclude researchers from collecting data on individual students however, as many studies included in this review outlined several benefits of evaluating individual students within a classroom context. For example, the effects calculated in a study by Dadakhodjaeva et al. (2019) indicate a much larger effect on academic engagement and disruptive behaviour for the whole class groups when compared with the three target individuals' data. In the study by Hoff and Ervin (2013), the opposite was true, with the effect sizes for individual student disruptive behaviour being larger than the effect sizes for the whole class. This indicates the importance of considering this type of data on a study by study basis. In the study by Dadakhodjaeva et al, baseline levels of academic engagement and disruptive behaviour among the whole class group were problematic and target students were selected based on their particularly problematic behaviour. In other words, an intervention was warranted for the whole class group, but certain individuals were chosen for individual monitoring. Hoff and Ervin observed less severe disruptive behaviour across three recruited classes at baseline but successfully brought three target students' behaviour more in line with their respective classes' behaviour. The outcomes of these two studies demonstrate the importance of study context and examining baseline behaviour when evaluating the effects of an intervention with an individual student versus a whole class group.

Through another moderator analysis, it was found that game-based interventions which were not specified as versions of the GBG by the authors were just as effective as games labelled as versions of the GBG. This demonstrates the promise for moving beyond the GBG in evaluating game-based behavioural interventions for classroom management, however it also must be noted that although many games were not specifically called the GBG, they were very similar to it. For example, the Quiet Classroom Game (Radley et al., 2016) is very similar to iterations of a positive GBG (or CBGG), in that students receive points throughout a class period for engaging in low levels of noise as a group (albeit, not in teams). The CW-FIT is also very similar to the CBGG in many ways, in that teams of students earn points for engaging in positive classroom behaviour. The previously discussed Fair Play Game was not considered a version of the GBG in this review, however the authors specify in their introduction that it was originally considered a modification of the GBG. This moderator analysis is therefore not as meaningful as looking at the actual mechanisms of the game-based intervention under investigation.

A final moderator analysis looked at whether games employing positive reinforcement in the forms of points awarded for appropriate behaviour (e.g., CW-FIT, CBGG), were more or less effective than games employing fouls or other mechanisms (e.g., traditional GBG). Positive games were not more or less effective in targeting disruptive or positive behaviours than games employing fouls. This is perhaps not surprising, given that even though the GBG involves positive punishment in the form of marks, students are rewarded for engaging in low rates of disruptive behaviour, which constitutes differential reinforcement of low rates of behaviour (DRL). This is well-established as an effective means of reducing target behaviour (Cooper et al., 2007). The GBG, although perhaps not desirable in modern classrooms based on its focus on rule-breaking and fouls, still maintains a large evidence base for its effectiveness over several decades.

Limitations

The findings of the current review must be considered in light of certain limitations. Firstly, studies were included in this review based on whether or not they were termed a ‘game’ by the researchers describing the intervention. It is possible that interventions exist which mimic game-based interventions in all but their title and therefore some studies may not have been included despite reporting on an intervention very similar to those included here. However, limits to inclusion/exclusion must be applied stringently in a systematic literature review and the decision was made based on ensuring clarity across reviewers. Second, while some effort was made to identify grey literature by including doctoral dissertations in the synthesis, other forms of grey literature were not included. Third, this synthesis was limited to mainstream/general education classrooms only, limiting the generalisability of results to special educational settings, despite the fact that game-based interventions have been applied in these settings with success (e.g., Flower et al., 2014; Groves & Austin, 2017). Future research may consider a review of game-based interventions in special educational and resource settings. Fourth, only studies meeting the WWC standards with or without reservations (WWC, 2017) were included in the study coding and meta-analysis stages of this review. Although this was incorporated to ensure studies included were of sound design, it led to exclusion of 19 studies at the design evaluation phase, reducing the body of literature included in the final analyses. This meant that many older studies were excluded, given poor reporting of IOA practices. This limiting of studies for meta-analysis and reducing the study pool reflects the stringent nature of a systematic review.

Implications for Research and Practice

The current review raises several issues which may be addressed in future empirical research and reviews and some of these issues will subsequently be addressed throughout this thesis. First, researchers may consider the range of game-based interventions available for empirical evaluation and look at the specific elements of these games that could be altered. It was evident in this review that many of the games evaluated were the GBG or versions of it, with the various versions trialled reflecting novel applications. The provision of points instead of marks (e.g., Wright & McCurdy 2012), the use of response-cost (Tanol et al., 2010) and the incorporation of modern, student friendly technology such as Class Dojo (Ford, 2017; Lynne et al., 2017) are just some ways in which the GBG has been extended in recent years. Some variations have been incorporated to align more closely with positive behavioural practices in schools. Less frequently mentioned, is altering a game-based intervention to save the teacher time in class or to provide the teacher with a sense of autonomy in implementing an intervention. It has been established that a teacher's sense of agency and freedom is important in their uptake of research-based strategies (Joram et al., 2020) and therefore it is important to identify features of interventions which are malleable. The recommendation to explore procedural variations of the GBG in terms of efficiency, efficacy and acceptability has been put forward in a recent narrative review by Joslyn et al., (2019), and the current review supports that recommendation for game-based interventions at a more general level. Considering these implications, the current thesis will progress to evaluate the most prevalent variation of the GBG, the CBGG, in Irish populations. The review has further established the need for general evaluation of the CBGG as distinct from the GBG, given its differing underlying principles. That is, the GBG combines DRL with positive punishment, whereas the CBGG incorporates positive reinforcement only. Although the CBGG represents a variation of the GBG in its own right, further procedural alterations to save teachers time and establish areas where teacher autonomy is possible is a useful avenue for research. The current thesis will therefore focus on the CBGG and its potential for variation while still maintaining effectiveness. Based on the findings of the systematic review, procedural variations including the provision of feedback, and the length of schedules between opportunities to earn points will be considered across different classroom contexts.

According to the review findings, the CBGG has not yet been evaluated in Ireland. As outlined in Chapter 1, there is a clear need for research on classroom management in Ireland, particularly practices which maintain a positive focus. The purely positive focus of the CBGG (as opposed to the GBG which incorporates positive punishment) makes it a suitable intervention for examination, given the current Irish educational climate's focus on positive behaviour supports as a first port of call where possible (National Council for Special Education 2020; National Educational Welfare Board, 2008). Furthermore, another interesting issue raised in this review which may influence future research, is that in reporting on gender in the studies included, most reported a mix of males and females in the class setting. In cases where this was not reported, it was sometimes unclear whether the intervention was applied in a mixed-sex or a single-sex setting. In other studies where target students' behaviour was monitored, gender of the target students was reported, but the gender make-up of the class was not (e.g., Tanol et al., 2010). Bowman-Perrott et al., (2016) specified the need for investigation of the potential mediating effect of gender in GBG research, given males demonstrate higher rates of problem behaviour than females (McIntosh et al., 2013). Most of the research on game-based interventions was conducted in the USA, where single-sex school settings are uncommon. This points to a need for research in cultures where single-sex education is more common, such as Ireland.

There were several game-based interventions identified in this review that have only been applied once in the literature. For example, the novel Quiet Classroom Game (Radley et al., 2016) and the Keep Busy and Carry On game (Hine et al., 2015) have the potential to be trialled with other populations. The game-based interventions here were diverse in many ways however, it was evident that most incorporated one key feature, that is, the interdependent group contingency. It is possible therefore, that small modifications to game procedures which do not alter the group contingency element could enhance teacher buy-in and save time. It is not within the scope of the current thesis to evaluate several types of game-based intervention in Irish classrooms, but the current review has served to identify other directions for future research.

In planning this review, the term gamification was explored as a potentially important modern term which could be used to describe game-based interventions applied in the classroom management literature. It was included as a search term when devising the search strategy and based on a narrative review of the literature, the term seemed to fit well with these intervention types. However, upon conducting the review, it seems gamification has not been a widely adopted term in this realm, despite recent discussions on the intersections between gamification and behaviour analysis (Linehan et al., 2015; Morford et al., 2014). Thus far, the term gamification has not been adopted in this thesis and 'game-based interventions' has been used instead to describe the types of intervention under study. Given that it is evident that 'gamified intervention' or 'gamification' have not been adopted as a terms in behaviour analytic literature, 'game-based intervention' will continue to be used in this thesis to maintain cohesiveness with the already large pool of literature.

For practical purposes, this review has succeeded in specifying the range of game-based interventions available to teachers of mainstream classes, and outlined their effectiveness overall and in relation to each other. Teachers may choose to adapt these games in their classroom settings, with the knowledge that game-based interventions can be effective across primary and second-level settings. This recommendation comes with the cautionary note that implementing any classroom management intervention and calling it a 'game' is not enough to ensure implementation of a research or evidence-based practice and thereby influence classroom behaviour. This review has found that the group contingency, particularly the interdependent group contingency, is a common key feature across most classroom management games. Practitioners may therefore build their skills in implementing group contingencies and present these contingencies as 'games' to potentially enhance buy-in from students. This finding also reiterates an important avenue for future research, which may consider whether 'game-based' group contingencies (i.e., those framed as games when presenting them to students) are more, less or similarly effective as those not framed as games.

Conclusion

The review and meta-analysis presented in the current chapter set out to identify behavioural game-based interventions in the classroom management literature, identify the characteristics of the student populations taking part in these studies and the key features of the games under investigation, and to assert whether game-based interventions have been effective in targeting positive and disruptive classroom behaviours in mainstream school settings. It is evident that game-based interventions have been relatively common in the literature and that most high-quality evaluations have taken place in the last 10 years. They have been effective in targeting classroom behaviour, however it is evident that group contingencies more often than not form a key feature of the game, which may be the crucial component to their success. Some clear lines of inquiry have been identified for future research, including investigation of procedural variations in game implementation, consideration of the gender make-up of the population under study and further probing the game-based intervention and group contingency link.

Chapter Three: Comparing Two Different Versions of the Caught Being Good Game with a Mainstream Adolescent Student Population[1]

[1] The first study presented in this chapter has been published in peer-reviewed format: 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

Introduction

The GBG has been identified in reviews, including the systematic review in Chapter 2 of the current thesis, as one example of an interdependent group contingency intervention which has successfully targeted engagement and disruption among students attending mainstream classrooms (Bowman-Perrott et al., 2016; Chapter 2). As previously described, this classic classroom management intervention typically involves the division of a class group into teams and the subsequent provision of marks to teams on which a member breaks a class rule. Teams remaining under a certain criterion of marks at the end of the game are eligible to receive a prize. The GBG and its variants, are the most popular game-based classroom management interventions assessed in the literature (Chapter 2). One of the common variants of the game, is a positive version, whereby students earn points for rule following rather than marks for rule breaking. This positive modification (often termed the CBGG) was discussed in Chapter 1 and identified in Chapter 2 as a promising game-based intervention. This chapter will focus on an under-studied population in CBGG research by applying the game with an adolescent student population. It will also deal with one specific procedural variation in examining whether visual feedback is a necessary component.

Positive Versions of the GBG

There was a paucity of research on positive versions of the GBG throughout the late 80s, 90s and early 2000s, with few high-quality studies published on the topic. This is evidenced in the review in Chapter 2, where three studies evaluated a positive version of the GBG alone (Ford, 2017; Hernan et al., 2019; Lynne et al., 2017) and two evaluated the CBGG and GBG together (Fallon et al., 2018; Wright & McCurdy, 2012). One further study evaluated the GBG-response cost in comparison with a positive version of the GBG (GBG-reinforcement; Tanol et al., 2010), and one evaluated a modified, positive GBG (albeit with a response-cost component) with a physical education class (Patrick et al., 1998). All but one of these studies were published since 2010, with studies implemented before that not meeting the WWC standards (e.g., Robertshaw and Hiebert, 1973) or otherwise not meeting the inclusion criteria for the review (e.g., Fishbein & Wasik, 1981). The emergence of high-quality evaluations of the CBGG in recent years has coincided with a move towards more positive behaviour strategies in schools generally. This has been evident in Ireland with recommendations by Tusla (the Child and Family Agency) to focus on positive behaviour in development of school behaviour policies (National Educational Welfare Board, 2008). It therefore makes sense to evaluate the CBGG with Irish school populations, to assert whether the intervention is effective. Up to this point in the current thesis, various terminology has been used to describe positive versions of the GBG, however from this point forward, the GBG will refer to any implementation of the game as it was originally put forward (i.e., marks against teams where rule-breaking occurs) and the CBGG will refer to any implementation of the GBG as a positive behaviour intervention (i.e., points for teams where students are following the rules).

The GBG and CBGG with Adolescents

There is a scarcity of research on the implementation of the GBG or CBGG with second-level students, that is, students from the ages of approximately 12 years and older. This is despite the fact that, any studies which have been carried out with this age group have produced promising results. Evidence from systematic reviews in the area suggest that group contingency and game-based interventions are just as effective with secondary school students as with primary school students (Bowman-Perrott et al., 2016; Chapter 2; Maggin et al., 2017). For example, the GBG was trialled with a ninth-grade history class and was very effective in reducing three sub-types of disruptive behaviour (Kleinman & Saigh, 2011). Mitchell et al., (2015) also examined the GBG with three high school classes and found that it could produce a reduction in disruptive behaviour in this age cohort. Despite these studies providing promising evidence, few other studies have subsequently examined the CBGG with adolescent students. One exception by Conklin et al., (2017) evaluated a group contingency intervention which was similar to the CBGG as part of the larger Class-wide Function-related Intervention Teams (CW-FIT) classroom management programme. The CW-FIT was effective in targeting on-task behaviour and compliance across two seventh-grade groups. Ford (2017) examined the CBGG in conjunction with point-awarding technology, ClassDojo. This is an interactive platform which allows teachers to award and display individual or team points on an interactive whiteboard (ClassDojo, 2019). The CBGG with the incorporated ClassDojo technology was effective in targeting academic engagement and disruptive behaviour in four high school classes. Despite this promising evidence, there are few other studies examining the CBGG with this age cohort, and this was evidenced clearly in Chapter 2. This represents a stark gap in the literature for evaluation of the CBGG with adolescent students. Although Chapter 2 identified several game-based interventions for potential evaluation with an adolescent population, the CBGG emerged as the most relevant avenue, given its positive focus and relative ease of implementation. One of the main aims of the studies presented in this chapter was therefore to evaluate the CBGG with an adolescent population.

Procedural Variations of the CBGG; Is Visual Feedback Important?

A procedural variation in recent GBG and CBGG investigations has been the provision of feedback. Some studies have opted to have the intervention agent (usually the class teacher) record marks or points in real-time on a display such as an interactive whiteboard (e.g., Lynne et al., 2017), while some have had the implementer record points discretely and only announce them when the game was over (e.g., Wahl et al., 2016). This procedural variation has been examined recently in the GBG literature, however the work has been largely inconclusive and there is much more to learn. Early evaluations of feedback delivery during the GBG were conducted by Medland and Stachnik (1972) and Harris and Sherman (1973). Medland and Stachnik (1972) put the GBG in place with two groups of fifth-grade students. Feedback was provided to teams in the form of a light system which was operated by observers and visible to the students. The light was changed from green to red for 30 s contingent on a member of the team breaking a rule. The researcher withdrew the game after a phase where the full GBG-package had been in place. After the withdrawal phase, the game was reinstated with rules only, and then with rules and lights only. Harris and Sherman (1973) conducted a similar study but evaluated feedback in the form of visible hatch marks rather than lights. Disruptive behaviour remained low when the GBG was put in place with no visible hatch marks, therefore the authors concluded that they were not an essential component. Despite providing interesting findings, both studies included the presentation of the full GBG-package before manipulation of the feedback elements. Therefore, it is impossible to delineate the potential crossover effects which may have occurred in ensuing phases of the studies.

Two more recent studies evaluating the feedback component of the GBG with preschool students were those by Foley et al. (2019) and Wiskow et al. (2019). Foley et al. (2019) examined several components of the GBG with a preschool class, one of which was feedback. Rather than first presenting the whole GBG-package, components were added one by one across several phases, culminating in the whole GBG-package. These components were rules, a criterion of marks, feedback (in the form of visible marks) and reinforcement (both non-contingent and contingent). The authors found that the GBG-package was needed to see a meaningful change in disruptive behaviour, however after the students had been exposed to the package, a version comprising rules, feedback, a criterion and non-contingent reinforcement could produce similar effects in keeping disruptive behaviour low. One issue with this study is that although it examined components of the GBG, it did so in a specific order, and did not examine whether the game was effective as a package but omitting the feedback component only. Wiskow et al. (2019) looked at feedback as an independent component during the GBG. Feedback was examined across four conditions: no feedback, visual feedback only, vocal feedback only and visual + vocal feedback. Vocal feedback only and visual + vocal feedback, were the most effective conditions in reducing the average number of disruptions per minute during the GBG. Despite the high-quality investigation, there are some limitations to this study which indicate many avenues for future research. For example, the experimenter implemented the game rather than the teacher. If students were attending to the class teacher, their attention may not have been drawn to the marks being given during the visual only condition. The authors noted that the participants, being preschoolers, may not have had the prerequisite skills to fully comprehend the delivery of marks during the game. For example, they may not have had the skill of counting to ten and understanding the concepts of more than/less than. As well as this, the game implemented during both previously mentioned studies was the GBG, and analysis of feedback during the CBGG has not yet taken place.

Research Questions

A number of research questions are proposed based on the current state of the literature around the CBGG with adolescents:

1. Is the CBGG an effective classroom management intervention in targeting academic engagement and disruptive behaviour in an ‘at-risk’, mainstream population of Irish adolescent students?
2. Is there a difference in effectiveness when the game is played with delayed feedback (CBGG-d) versus immediate feedback (CBGG-i)?
3. Does the implementing class teacher find the CBGG to be an acceptable intervention and do they prefer the CBGG-d or the CBGG-i?
4. Do part-taking students find the CBGG to be an acceptable intervention and do they prefer the CBGG-d or the CBGG-i?

STUDY 1

Method

Recruitment, Participants and Setting

A secondary school in Dublin was identified for participation in a group contingency project on the ‘STAR league’ (a group contingency intervention which the current researcher planned to run with first-year classes). This study specifically aimed to evaluate the CBGG with students deemed ‘at-risk’ of disadvantage. This was conceptualised based on the school’s designation of disadvantaged status. The school maintained DEIS status (Delivering Equality of Opportunity in Schools) according to the Department of Education and Skills. DEIS status is granted to secondary schools based on a range of criteria including, number of medical card holders, Junior and Leaving certificate retention rates, rate of students eligible for free books and Junior certificate results. Other variables considered include the number of non-national students, students from the travelling community and students with special needs (Weir, 2006). DEIS status has been used in previous research as a proxy for school social mix (e.g., McCoy et al., 2014; Quail & Smyth, 2014). This DEIS status designation was known to the researcher before approaching the school (this information is publicly available) and the school was approached because the students attending may be considered ‘at-risk’ for disadvantage. The principal was approached with an invitation to take part in the study. A representative from the school board of management read about the study in a plain language statement (PLS) and signed a consent form indicating the school’s willingness to engage in research (Appendix D1; D2). Parents and students read about the research and completed consent and assent forms (Appendix D3-D5).

A first-year class taught by Ms. Allen[2] were recommended for the project by the first year year-head. Upon approaching Ms. Allen and explaining the project, she expressed interest in taking part, however she was concerned about the complexity of the STAR league intervention. It was therefore decided to simplify the project substantially while keeping within the remit of the project for which the board of management and parents had provided consent. Thus, the current study on the CBGG was developed to be applied with Ms. Allen’s

[2] All teacher and student participants are assigned pseudonyms in this thesis, to ensure protection of participant identity.

mathematics class. Ms. Allen read about the study in detail and completed a consent form before data were collected during her class (Appendix D6; D7).

Ms. Allen's class was a first-year mathematics class. In an Irish school setting, first-year is students' ninth year of formal schooling and represents a transition to secondary school from primary school. Twenty-one parents/guardians and students returned consent and assent forms and therefore took part in the study (10 male, 10 female, 1 not reported). All 21 students were of Irish descent. The participants had a mean age of 12.6 years (age range = 12-14 years). Ms. Allen took the classes' mathematics lesson four times per week. She was a female mathematics teacher and was 33 years old, with five years of teaching experience. She had spent all of those five years teaching in the school in which this study took place. Ms. Allen had not used contingencies like the CBGG previously.

Materials

Materials needed for the game included two A3 laminated copies of the class rules, a vibrating timer app for Android/Apple (Tabata Timer: Interval Timer Workout Timer HIIT; Sharafan, 2018) which was installed on the teacher's Android smartphone, team leader boards to allow for recording of the team scores (laminated A3 pages for the white board; Appendix E) and reinforcers/prizes. A preference assessment was conducted to assert which prizes students may find most desirable (Appendix F). A list of prizes was devised with the help of the class teacher and students rated them by preference from one-six. The top three prizes identified from the preference assessment were school cinema passes, school shop tokens and 'good' journal notes. Other prizes which received high numbers of votes from some students included stationery. The teacher was provided with a protocol/game outline (Appendix G) and a simplified checklist (Appendix H) which were designed to help the teacher stay on track with the game throughout a class period. Note that the game protocol was developed before it was decided that prizes would only be given out weekly and before it was decided that a known criterion would be used and therefore the simplified checklist takes this into account. This same checklist was used by the primary observer (i.e., the student researcher) to carry out treatment integrity checks. The checklist was developed based on previous studies on the implementation of the CBGG in classroom contexts (e.g., Ford, 2017; Lynne et al., 2017; Wahl et al., 2016; Wright & McCurdy, 2012). Data collectors used a data collection sheet (Appendix I) and a pen to collect the data for this study. Intervals were signalled to the observers by a track saved to a smartphone through earphones and an earphone splitter was used on days where two observers were present.

Operational Definitions of Target Behaviours/Dependent Variables

According to Cooper et al., (2007), the operational definition of a target behaviour should be objective, refer to only observable behaviour, and be clear in terms of what is included and excluded in the definition. These guidelines were adhered to when compiling operational definitions of the target behaviours for this study, and guidance was obtained by referring to previous studies on engagement and disruption in the classroom (Mitchell et al., 2015; Radley et al., 2016; Wahl et al., 2016; Wright & McCurdy, 2012).

Preliminary observations of the class group were conducted to assert which behaviours were most problematic during the mathematics class period. These preliminary observations also ensured that there was minimal student reactivity when data collection commenced as both the teacher and the student group would habituate to the observer's presence. It was noted that many students did not engage fully with the material for much of the class time. Students also engaged in disruptive behaviours, such as playing with objects such as rulers and pens, and inappropriate vocalisations. The target behaviours identified and measured throughout the study, therefore, included student engagement (passive and active) and student disruption (verbal and motor). This was broadly in keeping with Wahl and colleagues (2016) definitions, however definitions of behaviours did differ in some respects. The definitions of these target behaviours follow.

Academically Engaged Behaviour (AEB)

AEB was the first target behaviour measured. This target behaviour was measured across two categories: active engagement and passive engagement. However, these were considered together as the composite variable, AEB. Active engagement was defined as follows:

> *"the student is actively engaged in the academic task assigned to them by the teacher and/or special needs assistant (SNA). This includes reading aloud, writing, copying from the board when instructed to do so, talking to a teacher or SNA about the task at hand or talking to a peer about the task when this has been permitted (i.e., sometimes students will be asked to work by themselves and sometimes correspondence with a peer may be allowed)".*

Passive engagement was defined as follows:

"the student is oriented towards the academic activity but is not actively engaged or engaging in any of the defined disruptive behaviours (i.e., the student's gaze is towards the relevant academic material). This includes looking at the board, looking at and listening to the teacher, SNA or peer or reading to oneself/silently".

Disruptive Behaviour (DB)

DB was the second target variable measured in this group. Different classes of disruptive behaviour were evident, however the predominant behaviours were inappropriate vocalisations and engaging in motor behaviour which was not related to the task at hand (e.g., manipulating objects in a manner incompatible with the ongoing task and placing of the head down on the desk). Disruption was therefore measured across two categories: off task verbal behaviour, and off task motor behaviour (similar to Wahl et al., 2016). Similarly to AEB, rather than coding the categories of disruption separately, both of these categories of behaviour were coded as DB. Off task verbal behaviour was defined as follows:

"the student is engaged in a vocalisation which was not authorised by the teacher, including talking/whispering to a peer, shouting, calling out, humming, singing or whistling. A student is not considered to be off task verbal if talking to the teacher or SNA about the assigned work. A student is also not considered to be off task verbal if talking to a peer when this has been permitted (see definition for active engagement). If a student is talking to a teacher or SNA due to their engagement in disruptive behaviour (i.e. they are being reprimanded at the time of recording) this can be classified as off task verbal, as the reprimand is disrupting their learning and has occurred directly because of some teacher observed disruption".

Off task motor behaviour was defined as follows:

"the student is engaging in movement not related to the assigned academic task for >3 seconds during the interval. This includes being out of seat without permission, turning around in their chair, sitting with their head on

the desk, playing with objects such as stationery or paper or swinging on two legs of their chair. Playing with objects can be defined as a student using one or two hands to manipulate an object, e.g., tap a pencil against the table, in a manner which is not compatible with the task at hand and which has not been authorised by the teacher/SNA".

Observation Procedures and Data Collection

Data were collected up to four times per week by the primary observer during mathematics classes. A second observer collected data alongside the primary observer on occasion to facilitate interobserver agreement data. This process will be outlined in further detail in the next section. Although the game took place during a 40 min class period, data collection sessions lasted between 10 and 20 min, depending on the observer's and teacher's schedules, with the majority of sessions lasting 20 min. Data were collected by individual-fixed partial interval recording and momentary time sampling. Observers did not interact with students while in the classroom collecting data. DB data were collected via partial interval recording and AEB data were collected via momentary time sampling (Cooper et al., 2007). When observers were signalled to begin observing a student, they would watch the same student for 15 s, and record the interval as either containing DB or not containing DB (partial interval recording). When the end of the interval was signalled, the observers would look at the same student and determine whether the child was engaged or not at that moment (momentary time sampling). A different student was observed every 15 s in a systematic and fixed manner, and each student was observed several times in one class period. These observation methods allowed for a percentage of intervals in which the target behaviour was present to be calculated. This could subsequently be graphed for analysis. For example, in a 20 min observation period, there were 80 intervals in which students were observed. If 40 of those intervals contained DB, then DB occurred during 50% of intervals. In a recent evaluation of observation techniques for group behaviour in the classroom, Briesch et al. (2015) recommended observing a different student every 15 s in order to get the most accurate representation of academically engaged classroom behaviour. Although observing students in a random rather than a fixed order was found to be slightly more accurate and representative of actual behaviour, Briesch et al. recognised that this may not always be possible or feasible where second observers become involved in data collection. Therefore, to facilitate interobserver agreement data in the current study, an individual-fixed method was chosen.

Data collection took place on consecutive days, allowing for scheduled school breaks (mid-term break, Easter break) and occasional breaks in data collection due to planned (e.g., parent teacher meetings) or unforeseen circumstances (e.g., teacher off sick, Storm Emma snowfall leading to school closures).

Observer Training and Interobserver Agreement (IOA)

IOA measures were conducted by trained undergraduate and graduate psychology students throughout the observation periods and these individuals are referred to as ‘second observers’. All second observers underwent a training session with the primary observer whereby they were given copies of the operational definitions and shown how to engage in momentary time sampling and partial interval recording procedures. They were given a chance to ask questions during this training session and again before they presented to collect data. Second observers had the primary observer’s email address and could contact her with questions at any time. Second observers were blinded to the study purpose and not told what type of phase (baseline or intervention) they would be collecting data in ahead of presenting to collect data.

IOA data were collected during 25.64% of total observation sessions and were conducted in line with the WWC design criteria for single case research designs (WWC, 2017; 2020). Specifically, as outlined by the WWC, IOA observations took place at least once per phase and on at least 20% of observation sessions in baseline and intervention phases. IOA rates were calculated by placing the number of intervals in which observers agreed, over the number of agreements plus disagreements, and multiplying by 100 to obtain a percentage. The goal was to keep IOA levels above a mean of 80%, a guideline put forth by the WWC. If IOA was below 80% with any second observer, the primary observer met with this observer to retrain them before they would present to collect data again. This occurred during less than 40% of sessions and mean IOA remained over 80% on average overall for each outcome variable. Mean IOA for DB was 90.56% (range = 81.94-100%) and mean IOA for AEB was 84.46% (range = 72.15-100%).

Social Validity

Ms. Allen completed the Intervention Rating Profile-15 (Appendix J) following the final day of data collection (IRP-15; Martens et al., 1985). This rating profile is made up of 15 statements which are compiled to assess a teacher’s acceptability rating of an intervention under investigation. Items from the IRP-15 were modified to reflect the application of an intervention to a group of students (e.g., ‘child’ changed to ‘students’) and to the present/past tense. For example, “This would be an acceptable intervention for the child’s problem behaviour” was modified to “This was an acceptable intervention for the problem behaviour(s)”. Similar changes have been employed in other studies on the GBG (e.g., Mitchell et al., 2015). All items were phrased positively, and items were rated from 1 (strongly disagree) to 6 (strongly agree). The minimum score, therefore, was 15 and the maximum, 90. One additional item was added, asking the teacher whether she preferred the CBGG-d or the CBGG-i.

The students were given a modified version of the Children’s Intervention Rating Profile (CIRP; Appendix K; Mitchell et al., 2015; Witt & Elliott, 1985). The modified CIRP is a social validity measure with eight items such as “Did you like participating in the Game?” to which students simply answer ‘yes’ or ‘no’. The modifications to the original CIRP were adapted from Mitchell and colleagues (2015) and involved changing the tense of items from present to past tense, using the term ‘student’ instead of ‘child’ to suit the adolescent population and the addition of one item on rewards used. A yes/no format was used also instead of a Likert scale. In the current study, the final two items were changed from positive wording to negative wording to enhance clarity. The highest rating a student could give the intervention was eight (eight positive endorsements). An additional question was added asking students whether they preferred the CBGG-d or the CBGG-i.

Experimental Design

A reversal/withdrawal design, with phases ABACABAC, was used in this study. This is a variation on the standard reversal/withdrawal design. The variation in design here involved the addition of a C phase, in which the same intervention was applied but with a significant variation in procedure (i.e., feedback administered publicly). In the current study, phase A was a baseline phase where regular classroom management procedures used by the teacher were in place. Phase B refers to the CBGG-d and phase C refers to the CBGG-i. Although some studies have evaluated two versions of the GBG using less phases (e.g., ABCBC; Wahl et al., 2016), the current study incorporated withdrawal phases between each intervention phase to further enhance the strength of the study design. This method maintains consistency with the WWC standards for single case research designs (WWC 2017; 2020). These standards state that for a reversal/withdrawal design to meet single-case design standards there must be at least three attempts to demonstrate an intervention effect. The current study incorporated withdrawals between each B and C phase, maintaining the WWC design standards, while allowing for comparison between the two game versions.

Phase sequence and changes were determined a-priori. Randomisation of phases was not desirable as a) it was important that the CBGG-d was the first to be presented to participants (i.e., that there was a component withheld) and b) for the teacher, with her busy schedule, it was desirable that she could pre-plan for many weeks ahead. The game in place always finished on a pre-set day (Friday). This allowed for the provision of prizes at the end of the week and allowed students to be made aware of the finishing day and points criterion, and work towards this goal. Phase lengths were generally predetermined, and where possible the game would be played across one week, however if the case arose where data were not stable towards the end of a phase, it could be extended into the following week. This decision was always made in conjunction with the teacher and no phase change decisions were altered without her input. An example of when this occurred is evident in the graphed data in the results section (Figure 3.2.). At data point 9, it had been planned to return to baseline, however in conjunction with the teacher, it was decided that the game would be put in place again the following week, given the instability in the data collected in the run up to data point 9.

Procedure

Baseline

The baseline phase consisted of regular classroom instruction and lasted for five sessions. The teacher did not put any CBGG contingencies in place and engaged in her regular disciplinary actions. No specific reinforcement contingency for well-behaved and engaged students was in place. Disciplinary action for DB was taken in the form of verbal warnings, giving journal notes and sending consistently disruptive students out of class for a period, to their form teacher (head teacher for their class group) or to their year head (head teacher for their year group).

Teacher Training

After baseline measures were taken, the primary observer met with the class teacher to conduct training in the CBGG methods during one free class period (35 min). During training, the teacher was provided with an outline for intervention implementation which described both conditions. This included, introduction of the game, announcement of game rules, announcing that team names will/will not be written on a daily leaderboard (depending on game version) and stating that points will be recorded at random times throughout the class publicly on the board/privately in the teachers' notebook. The teacher and researcher together devised a set of rules to be used during the game. These were: "I will do my best to stay on task during class" and "I will allow my classmates to concentrate and learn". There was also a rule on having class materials required every day, however data were not collected on this as a target behaviour. The researcher showed the teacher how to set up and use the Tabata timer app and they together decided to set the intervals at 5 min to allow for minimal distraction to the teacher during class time, while still maintaining a relatively dense schedule of reinforcement. During training, the teacher expressed concern at administering prizes daily during the CBGG. The classes were only 40 min long and time was of the essence in trying to cover all class material. It was therefore decided that a weekly criterion be trialled, with students receiving prizes based on their points total at the end of a week. It had also been planned that a mystery criterion would be used during game implementation, however the teacher again expressed concerns with this and therefore, to enhance buy-in, a known criterion was used. A known criterion has been commonplace in GBG and CBGG research and the game has been effective with known criterions in place (e.g., Barrish et al., 1969; Ford, 2017; Kleinman & Saigh, 2011). The teacher and researcher discussed reasonable aims for points needed to obtain the prizes/reinforcers which would be available at the end of each game phase. These aims were adjusted throughout intervention phases based on the students' performance. Treatment integrity was monitored by the primary observer in the form of a checklist which both the observer and teacher had access to. This checklist was a condensed version of the game outline which was easier for the teacher to follow during the class period.

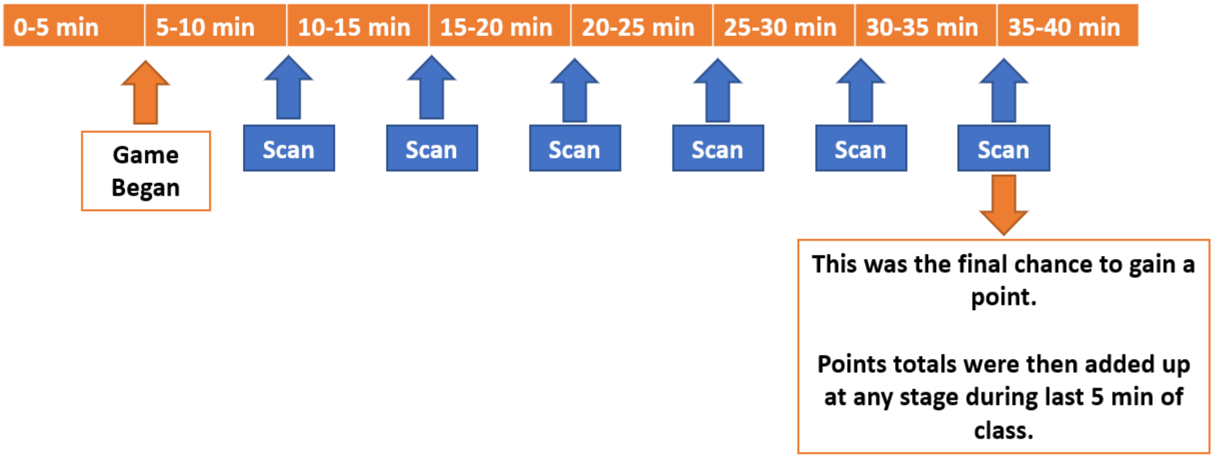
Intervention: Caught Being Good Game

Following baseline, the first game phase began with the teacher announcing to the class that a team-based game would be played and dividing the class into three teams. The teacher decided how many teams there would be and who would be on each team and based this on the layout of the classroom. The classroom was arranged into three columns of students sitting in pairs and therefore each column was a team. It is important to note however, that particularly disruptive students, as identified by the teacher, were dispersed evenly among the teams ensuring teams had an equal chance of winning. Students were given time on the first day of the CBGG to choose team names. Rules were posted in two locations in the classroom and students were reminded daily of the rules.

Upon announcing that the game would be played and reviewing the rules, the teacher announced the beginning of the game, the timer was started, and the teacher proceeded with her planned educational tasks for the day. The timer would vibrate at 5 min intervals, and upon hearing/feeling the phone vibrate, the teacher would award points to teams on which all members were following the rules. This meant that the timer would vibrate at six different stages throughout the 40 min class as the first and last 5 min were not considered part of the game (see Figure 3.1. for an example of the teacher's behaviour checking schedule). To ensure that the intervals were not exactly fixed and therefore predictable by students, the teacher was asked to award points at any stage from 0-60 s after the phone vibrated. A similar practice was applied by Ford (2017) and can be considered an alternative where a device set to variable intervals, such as a MotivAider, is not available. This practice may also be more practical for teachers who have access to a smartphone device but not to a MotivAider.

Figure 3.1.

A Visual Depiction of the Behaviour Checking Schedule Used by the Teacher During the CBGG



A points criterion was set based on how many days the game would be in place for and on how many points could possibly be earned within that time frame. For example, if the game was to be played for six days, the maximum points possible would be 36 (6 points x 6 days). The teams could earn an additional 'bonus point' if they presented to class with all of their materials, leaving the possible maximum number of points at 42 if the game was to be played over six days. The researcher and teacher would come to a decision before the game phase would begin on how many points the teams could feasibly hope to achieve. The goal was to make the criterion achievable yet high enough to encourage behaviour change, and varied from 50-75% of total available points throughout the study. Teams meeting the criterion by the end of a series of classes were eligible to gain access to the prize for the week. Friday was noted by the teacher as the most appropriate day to give out a weekly prize, given that it signified the end of a school week. Therefore, the series of games was always set up to end on a Friday. This prize and criterion were announced on the first day of gameplay so that students knew what they were working towards. If students struggled or failed to meet this criterion during the first game phase, the criterion could be subsequently reduced for future game phases. This was monitored consistently by the primary observer and teacher and decisions were made on a week to week basis. Similarly, if a team hadn't earned enough points by the beginning of Friday's class to be in with a chance of earning the weekly prize within the class period, smaller prizes were made available with a daily criterion put in place. For example, if the weekly criterion was 30 points and a team could earn 7 points per class, a team with only 20 points by the beginning of Friday's class could not hope to achieve 30 points. The teacher would then announce a smaller prize which could be earned by all teams earning three points on that day.

Feedback Phase (FB). As there had been no opportunity for the teacher to practice implementation during training, a brief FB phase was put in place to ensure treatment integrity was high. This phase involved the researcher giving immediate feedback to the teacher on their game implementation after the relevant class, until the teacher reached 80% treatment integrity. The teacher implemented the game with 70% integrity on the first day, so feedback was given. On the second day, integrity exceeded 80%, therefore the FB phase was terminated.

CBGG-d. The CBGG-d involved the implementation of the game as described. When the teacher scanned the room, she would record points privately in a notebook and students would only find out how many points their team had received at the end of class. At the end of class, the points were placed onto the weekly leader board so students could see how many points they had earned and how many more were needed in order to earn the top prize at the end of the week.

CBGG-i. The CBGG-i involved the points being displayed on the leader board under the relevant team name immediately after the teacher had scanned the room. At the end of class, these points were tallied and added to the overall weekly leader board so students could keep track of their progress towards the prize. The teacher had been encouraged during training to pair awarding of a point with a positive comment during this phase, however this was not a requirement on the treatment integrity checklist, nor were data collected on the teacher's rate of praise.

Treatment Integrity

The teacher checklist, including 10 steps for completion of the game, was given to the teacher, and she was asked to complete each step daily to ensure that the game was implemented correctly. She did not have to complete the checklist manually but had the option to do so and was asked to keep it on her desk to refer to. The primary observer completed the checklist during all observation sessions in intervention phases to obtain a measure of treatment integrity. No specific contingencies were in place if treatment integrity was low (<80%), however the primary observer consistently encouraged the teacher to use her checklist and implement all steps of the CBGG.

Treatment integrity in implementing each of the 10 intervention steps ranged from 30-100% (*M* = 77.6%) A breakdown of treatment integrity in each intervention phase is displayed in Table 3.1. There were four steps which the teacher most often missed when implementing the game. The first was reminding students of how many points they needed to obtain the prize. The last three steps of the game were also commonly missed. These involved announcing the game was finished, announcing and recording team points and reminding students of how many points were now needed to get a prize/announce the winners (if it was the last in a series of games). Steps most often missed all occurred at the end of the game period and missing the steps was usually due to time constraints at the end of class. Ms. Allen generally addressed these steps at the beginning of the following class and before starting the next game session.

Table 3.1.

Mean Rates and Range of Treatment Integrity during Intervention Phases for Ms. Allen

Phase	*M*	Range
Feedback	75%	70-80%
CBGG-d 1	85%	60-100%
CBGG-i 1	71%	30-100%
CBGG-d 2	78%	50-100%
CBGG-i 2	78%	70-100%

Data Analysis

Design Evaluation

The study design was evaluated using the WWC criteria for single-case research designs (WWC, 2017; 2020). A study can be found to meet the design standards with or without reservations, or not to meet standards based on these criteria. The criteria were the same as those applied in the systematic review presented in Chapter 2.

Visual Analysis

Visual analysis of the data was carried out in line with recommendations for evaluating single-case research (WWC, 2017). These procedures involved assessment of changes in level, trend and variability in data patterns between phases and across phases which are similar. It also involves looking at the immediacy of effect and extent to which data overlaps between phases. Consistency of data patterns across phases which are the same were also examined.

Effect Size

To supplement visual analyses, effect sizes were calculated using the Tau metric. Tau is an effect size metric based on Kendall’s Rank Correlation. In the current study it was calculated based on Tarlow’s (2017) recommendations, as was practiced in the review in Chapter 2. Tau values were calculated for each separate AB and AC phase contrast using the Baseline-corrected Tau calculator (Tarlow, 2016). Weighted mean effect sizes were calculated for both versions of the game (CBGG-d and CBGG-i) and for both outcome variables (AEB and DB). This was done by weighting effects for each phase transition by their inverse variances (Tarlow, 2017) and calculating a weighted mean effect size using these weights. This can be calculated in Winpepi (Abramson, 2011), as outlined in the Method section in Chapter 2. An effect size of .20 may be considered small, .20-.60 moderate, .60-.80 large and .80+ very large (Vannest & Ninci, 2015).

Ethical Considerations

The Psychological Society of Ireland (PSI) and the British Psychological Society (BPS) publish Codes of Ethics, available to members and non-members and which can be considered when planning, carrying out and writing up any research project (BPS, 2010; PSI, 2010). The DCU Research Ethics Committee (REC) also provides a set of guidelines and principles which must be adhered to in research processes conducted in the university by staff and students. All of the above guidelines were considered when developing an ethical protocol for this study and ethical approval was sought from DCU REC. The current study was approved by DCU REC before recruitment commenced (Appendix L). Ethical considerations in the current study will be discussed under the three ethical principles put forward by DCU REC.

Principle 1: Respect for Research Subjects and Participants

This principle was adhered to in study planning and recruitment, throughout data collection, in data storage and in reporting on data collected. All recruitment methods were pre-approved by DCU REC before study commencement. Protocols were followed to ensure participants, including teachers, parents and students, didn't feel they were obliged to participate (opt-in consent rather than opt-out). It was important that they had a full understanding as to what they may be asked to do, how data would be collected on their behaviour and that they were informed about the subsequent storage of this data. Risks identified were deemed to be not greater than those which teachers and adolescents may encounter in everyday life (i.e., classroom management practices including rewards and teams are common in classrooms and available to all teachers to use), however relevant supports were put in place for teachers and students who may have experienced adverse reactions, or simply had more questions about the nature of the research. Parents, students and participating teachers were provided with the student researcher's contact details, as well as contact details for the research supervisor and an independent party from DCU REC. A statement in the PLS noted that the researcher would address any additional questions concerned parties may have had.

All data collected was anonymised such that no names were attached to written data and the names of teachers and/or schools are not reported during any stage of the research process. Behavioural data was collected on a group basis so behaviour data could not be attributed to any one particular student. Teachers were advised that they may be identifiable in their school due to the small number of part-taking teachers within the school, however their names were never used during data collection or when writing up findings (pseudonyms were used). It is also important that any funding sources and conflicts of interest are made known to part-taking parties. In this case, there were no conflicts of interest and the student researcher's funding source was reported.

Principle 2: Fairness and Justice

Selection of participants should be fair and based on the research aims. The target population for this study were a group of secondary school students demonstrating issues around school-based behaviours (e.g., attendance, disruption, engagement) such that their learning time was potentially disrupted. The school principal and board of management approved researcher access to the school before a meeting was arranged to discuss the study. The first-year year-head suggested a teacher who may be interested in taking part. The year-head approached the relevant teacher and only when the teacher had expressed to the year-head that she would be interested in taking part, did the student researcher make further contact with her. The class targeted and their teacher had no prior engagement with the student researcher such that the researcher had no influence over them in terms of willingness to participate or subsequent behaviour change. The class teacher read about the study in a PLS and signed a consent form indicating her willingness to take part. Students whose parents had not returned consent forms could still engage in the classroom management procedure given the teacher had decided to adopt it, however data would not be collected on their behaviour as part of the group data collection.

Principle 3: Participant Well-being

All potential risks and harms were considered prior to recruitment and any risks identified were flagged by the researchers during the REC approval process. As previously mentioned, the risk identified was not greater than that which may be encountered in the everyday life of a student in secondary school. Teachers have a number of classroom management strategies which they may use in their classroom and it is highly likely that participants had previously experienced a) some type of reward-based contingency for good behaviour and b) a team/group-based consequence in the classroom. Contingencies were put in place for students who may be distressed. Cases of severe distress noted by the teacher or researcher, would result in the intervention being withdrawn until further notice and the relevant counselling/behaviour support services in the school being notified. The student researcher ensured she became familiar with student services in the school setting and made herself known to the first-year year head, so that if issues did arise, she had a point of contact relevant to the group. There were no apparent incidences of distress for teachers or students throughout this study.

Results

WWC Design Standards

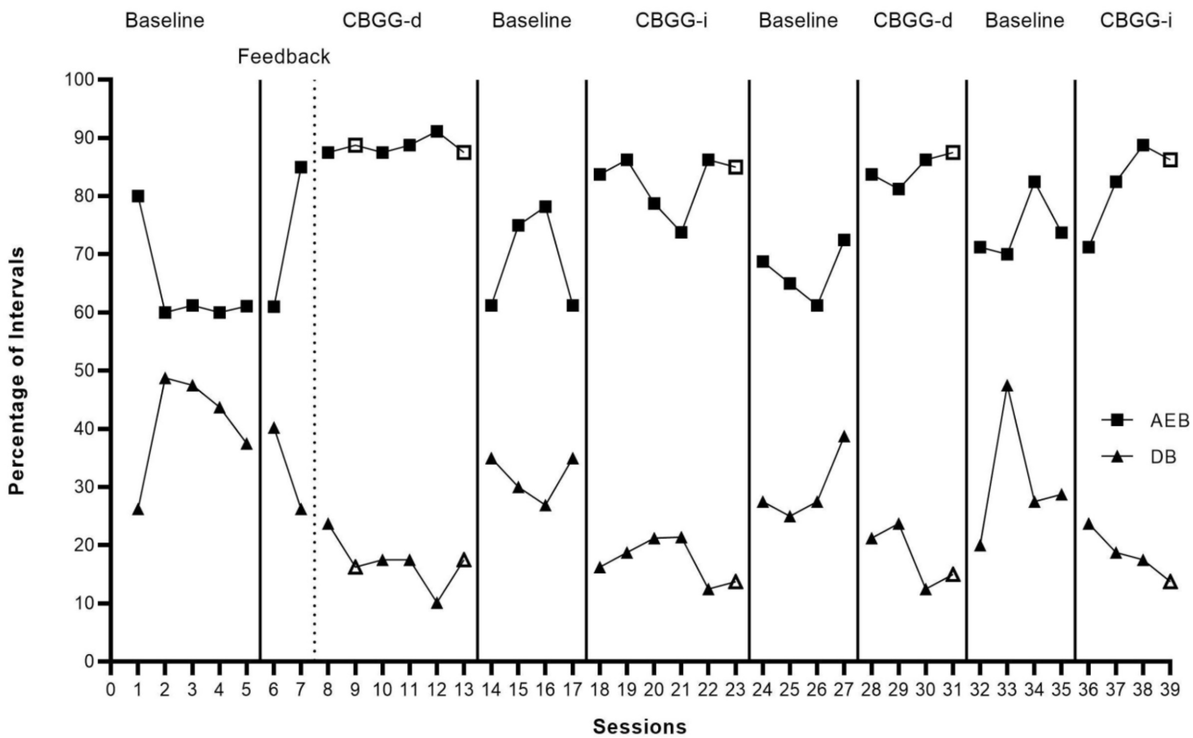
As per the WWC design standards (WWC, 2017; 2020), the independent variable (i.e., the implementation of the CBGG) was systematically manipulated throughout the study. IOA data were collected at least 20% of the time (overall and within each study condition), at least once per phase, and met minimum thresholds of 80%. Each phase within the study included at least three data points. Although five are preferable to meet the WWC standards without reservations, in the current study five data points were sometimes unreasonable to attain within one school week and within the timeframe of the study. The study employed a reversal/withdrawal design (ABACABAC) to ensure sufficient replication of each intervention condition. There were at least three attempts to demonstrate an intervention effect for each version of the intervention. Finally, the WWC states that the study should attempt to remain free from residual intervention effects when two interventions are being compared. This was counteracted in this study by applying the version of the game with delayed feedback first. This meant that students had not been subjected to all game features/components in the first intervention phase. The teacher remained the same throughout the whole school term and she delivered the intervention on each occasion. No other confounds were apparent throughout the process. The study therefore meets WWC standards with reservations because there are less than five data points in some phases.

Visual Analysis

Figure 3.2. displays percentages of intervals in which AEB and DB were present in Ms. Allen’s mathematics class.

Figure 3.2.

Percentage of Intervals with AEB and DB across Study Phases for Ms. Allen's Mathematics Class



Note. Open squares and triangles denote sessions where a weekly prize was available (i.e., a Friday).

Academically Engaged Behaviour

AEB occurred during a mean of 64.47% of intervals (range = 60-80%) during baseline sessions which is considered a low rate. Although the first data point suggests engagement was high, this quickly decreased and stabilised over the following four data points. During the FB phase, AEB remained low during the first data point, however when the teacher and students became familiar with the game, AEB increased immediately and substantially (*M* = 73%, range = 61-85%). This increase was sustained during the first CBGG-d phase. AEB was recorded during a mean of 88.52% of intervals (range= 87.5-91.13%) during this study phase. There was a substantial change in level of the data between the first two experimental phases and there was no overlap between the first CBGG-d phase and the initial baseline phase.

During the first withdrawal phase, AEB decreased immediately and substantially, occurring during a mean of 68.93% of intervals (range = 61.25-78.2%). There was no overlap between this withdrawal phase and the preceding intervention phase. The first CBGG-i phase corresponded with an immediate and substantial increase in AEB (*M* = 82.3%, range = 73.8-86.25%), however levels did not remain as stable as during the previous CBGG-d phase. There were two apparent decreases in AEB in the middle of the phase (points 20 and 21; Figure 3.2.), however only data point 21 overlaps with the preceding withdrawal phase.

There was an immediate and moderate decrease in AEB on introduction of the second withdrawal phase (*M* = 66.88%, range = 61.25-72.5%). On reimplementation of the CBGG-d, AEB increased again (*M* = 84.69%, range = 81.25-87.5%) and remained high and stable across the phase. This reflects a large change. There was no overlap between this phase and the preceding withdrawal phase.

During the final withdrawal phase, AEB occurred in 74.38% of intervals on average (range =70-82.5%) however during one session (observation session 34; Figure 3.2.) there was a stark increase in AEB which was uncharacteristic compared to the rest of the phase. The final phase, involving four sessions of the CBGG-i, initially did not appear to influence AEB. However, towards the end of the phase, AEB increased and the mean level was higher than any of the baseline phases (*M* = 82.19%, range = 71.25-88.75%).

Disruptive Behaviour

At baseline, DB occurred at a high rate (*M* = 40.75%, range = 26.25-48.75%). Behaviour did not improve immediately on the first day of CBGG-d implementation in the FB phase, however after one session, DB decreased substantially (*M* = 33.25%, range = 26.25-40.25%) and remained stable across the first CBGG-d phase (*M* = 17.11%, range = 10.13-23.75%).

During the first withdrawal phase, DB increased immediately and occurred during a mean of 31.73% of the intervals (range = 26.92-35%). There was no overlap with the previous intervention phase. DB decreased again during the first CBGG-i phase (*M* = 17.32%, range = 12.5-21.42%). Despite demonstrating a slight upward trend initially, DB remained relatively stable and low across the phase and did not overlap with the preceding withdrawal phase.

During the second withdrawal phase, DB increased immediately, remained stable and then increased further towards the end of the phase (*M* = 29.69%, range = 25-38.75%). During the second implementation of the CBGG-d, DB decreased immediately and was low and across the phase (*M* = 18.13%, range = 12.5-23.75%). There was no overlap with the preceding withdrawal phase.

In the final withdrawal phase, DB increased modestly initially, however the level remained higher than during intervention phases (*M* = 30.94%, range = 20-47.5%). When the CBGG-i was put in place for the final phase, an immediate decrease was not apparent for DB, but behaviour demonstrated a consistent decreasing trend throughout the phase (*M* = 18.44%, range = 13.75-23.75%).

Effect Size

Tau values for each phase contrast, intervention version and outcome variable are reported in Table 3.2. Weighted means are also presented for each version of the game and outcome variable.

Table 3.2.

Tau Effect Sizes Across Phase Changes for Ms. Allen's Mathematics Class

Phase Change	AEB	DB
Baseline to CBGG-d (FB phase not considered)	.775**	-.76**
Withdrawal to CBGG-i	.623*	-.74*
Withdrawal 2 to CBGG-d 2	.756*	-.77*
Withdrawal 3 to CBGG-i 2	.49	-.66
Weighted Mean CBGG-d	.77	-.76
Weighted Mean CBGG-i	.57	-.71

Note. Baseline corrections were not required for any of these calculations.

**p<.05, **p<.01*

Effect sizes were large (>.60) across every phase change, except for AEB in the transition from the final withdrawal phase to the CBGG-i (Tau = .49). This was a moderate effect size. Weighted mean effect sizes were large for each intervention and outcome, except for the effect of the CBGG-i on AEB (Tau = .57), which was a moderate effect size.

Social Validity

Teacher Rating

Ms. Allen completed the IRP-15 (Martens et al., 1985) based on the CBGG in general and was asked to denote a preference for one version or the other. Ms. Allen's score on the IRP-15 was 64 out of 90 (M = 4.27). She slightly agreed, agreed, or strongly agreed with all statements except for two with which she slightly disagreed. These were: "This intervention was reasonable for the problem behaviour(s) described" and "This intervention was a good way to handle the behaviour problem(s)". Ms. Allen denoted a preference for the CBGG-d, noting that it led to less disruption of class time. Some written qualitative feedback was also sought out from the teacher which clarified and built upon the questionnaire responses. She stated that although the intervention was beneficial for most of the students, that some students still needed extra behavioural supports and strategies. She also noted how it could be unfair that due to prizes being awarded on a Friday for overall weekly behaviour, students who had poor attendance over the week could still get prizes.

Student Rating

Eighteen students completed the modified CIRP (Mitchell et al., 2015; Witt & Elliott, 1985). The mean score across the students who completed the questionnaire was 6.17. Responses were generally positive, with students deciding the game was fair, that they liked the rewards and that they enjoyed participation. Eleven students said that they preferred the CBGG-i (61.11% of respondents) and six said that they preferred the CBGG-d (33.33% of respondents; one student did not report a preference). Two students left comments indicating that they felt sometimes a team member could ruin their chances at a prize (e.g. "I think you should give other people a chance with different groups"). One stated that they did not want to play the game and rated it 4/8. Finally, one student left a positive comment, simply stating that s/he liked playing the game.

Summary and Brief Discussion

The current study adds to the literature supporting the effectiveness of the CBGG as a classroom management intervention in a mainstream population (Wahl et al., 2016; Wright & McCurdy, 2012) and more specifically in a second-level population (Ford, 2017). The intervention was effective in two formats in targeting AEB and DB in a first-year group. It builds upon previous research by comparing two versions of the game; one with delayed visual feedback (the CBGG-d) and one with immediate visual feedback (CBGG-i).

Effectiveness of the CBGG with and without Visual Feedback

The CBGG-d and the CBGG-i had similar effects on AEB and DB in Ms. Allen's mathematics class. Mean rates of DB always decreased and mean rates of AEB always increased when the CBGG was in place compared to preceding baseline and withdrawal phases. There were also large mean effect sizes for the impact of the game on both behaviours. These findings are in line with previous research suggesting that group contingency interventions are effective in targeting challenging behaviour (Maggin et al., 2012; 2017) and also those suggesting that the GBG is an effective classroom management intervention (Bowman-Perrott et al., 2016). Although both versions of the game appeared effective across intervention phases, stronger intervention effects were observed in earlier phases compared to the final game phase (i.e., the second iteration of the CBGG-i). This may have been due to novelty effects with the introduction of the game in earlier game phases or the use of more desirable prizes in earlier game phases. The CBGG-d appeared to be particularly effective with the students, with large effect sizes for its impact on both AEB (Weighted Mean Tau = .77) and DB (Weighted Mean Tau = -.76). It is also evident in the graphed data that there were immediate and stable increases in AEB and immediate and stable reductions in DB when the CBGG-d was put in place. Again, this must be interpreted with caution, as this was the first version of the game to be introduced, and the prizes used, particularly the cinema pass used towards the end of the first phase of the CBGG-d, was particularly desirable for the students. This will be discussed further in the Limitations and Future Directions section later.

Social Validity of the CBGG with Teacher and Students

Both the teacher and students found the game acceptable, with the teacher rating it 64 overall on the IRP-15 (Martens et al., 1985) and the students giving it a mean rating of 6.17 on the modified CIRP (Mitchell et al., 2015; Witt & Elliott, 1985). These ratings are comparable to ratings by high school teachers and students in a recent GBG study (Mitchell et al., 2015). The part-taking teacher preferred the CBGG-d however most students preferred the CBGG-i. The teacher's preference is potentially due to the time saving involved with the CBGG-d. In future, the use of the Behaviour Intervention Rating scale (BIRS; Elliott & Von Brock Treuting, 1991) may be useful to assess the teacher's perception of game effectiveness and efficiency as well as acceptability. Student preference is potentially due to the generalised conditioned reinforcers in the form of points, in that they had likely experienced the provision of visual points feedback during sports and other game-based activities.

Limitations and Future Directions

Although the results of the current study are promising, some limitations must be considered. Firstly, only one class and teacher in one school setting were recruited. This greatly limits the generalisability of the results to other class groups. As well as this, as two interventions were being compared, recruitment of two class groups would have allowed for counterbalancing of intervention conditions (i.e., the application of an ACABACAB design with a different class group). The prizes/rewards used throughout the game varied and included school cinema passes, edibles, stationery, and school shop tokens. Some of these prizes were costly and the researcher provided the teacher with them. Had the game been applied without the researcher present, the teacher may not have had access to prizes which needed to be bought. School cinema passes were the most preferred prize but could only be used for one game phase (the latter half of the first CBGG-d phase). Although these passes did not have a monetary cost, students would have to miss three class periods to attend the school cinema, meaning they were time intensive. This type of prize is desired by the students but potentially not by teachers on a regular basis. Data were only collected on student behaviour and teacher behaviour was not considered. Anecdotally, the observers did not witness a high rate of teacher praise, even when the CBGG-i was in place, despite the teacher being encouraged to pair the awarding of a point with a positive comment. Conclusions cannot be drawn about this as data were not collected, therefore an interesting line of inquiry for future research may be to monitor teacher statements during the CBGG-d and the CBGG-i. Finally, data were analysed on a group-basis, meaning individual inferences about improvements in behaviour cannot be made.

Conclusion

The current study provides preliminary but important evidence for the use of the CBGG with mainstream Irish adolescent students. Furthermore, the population recruited were deemed ‘at-risk’ based on the school’s DEIS designation, targeting an important population. The game was effective both with and without immediate visual feedback and teachers and students rated the game positively. More research is needed in the area to enhance generalisability of findings. A study which counterbalances intervention conditions across two classes would also be a useful point for future research to control for order effects, and this was the purpose of Study 2.

STUDY 2

Introduction

Study 1 was a preliminary investigation into the CBGG with an at-risk adolescent population, while manipulating the provision of visual feedback during the game. The study had some limitations which have been outlined briefly, the main limitation being the recruitment of one class group which meant intervention conditions could not be counterbalanced and that results were limited in their generalisability. The primary purpose of Study 2 was therefore to expand upon Study 1, by replicating the intervention across two first-year mathematics classes, thereby using a more robust research design. Study 1 employed a reversal/withdrawal design, with one participant case (i.e., Ms. Allen's mathematics class). Although there were a sufficient number of phases and data points per phase so that the design met the WWC design standards with reservations (WWC, 2017; 2020), extending the study with an additional case with counterbalancing of conditions would provide more robust findings. Two first-year mathematics classes were recruited for participation in Study 2; an ABACABAC design was used in one class (Ms. Brady's class) and an ACABACABAB design in the other (Mr. Carroll's class), to counterbalance the order of implementation of each version of the game across classes.

A secondary aim of Study 2 was to monitor teacher statements during game implementation to investigate whether the CBGG, in either format, impacted rates of teacher praise or reprimands. It was anecdotally noted by observers during Study 1 that the class teacher did not seem to increase her rate of praise while the CBGG was in place and would often award points without making any reference to the positive and appropriate behaviour engaged in by the class. Blaze et al., (2014) recently found that an intervention which involved prompting teachers to provide praise produced increases in AEB and decreases in DB across four classes of adolescent students. The CBGG is a positive behaviour intervention in which the teacher has several opportunities to provide praise to students. Given the powerful impacts praise can have on student behaviour without an additional intervention, it is important to examine the potential additive effects it may have during the CBGG. Teacher statements have been examined in the context of the GBG and CBGG previously, with authors examining whether implementing the intervention led to changes in teacher rates of praise and/or reprimands. For example, during Lannie and McCurdy's (2007) evaluation of the GBG with a first-grade class, teacher praise did not increase while the GBG was in place, despite decreases in disruptive behaviour being observed. Negative/neutral statements did decrease slightly as disruptive behaviour decreased. In another study comparing the GBG-response cost with the CBGG, Tanol et al. (2010) observed a slight increase in praise during CBGG phases. This corresponded with decreases in rule violations by students. In a more recent example, Wahl et al. (2016) compared the effects of the GBG and CBGG on student behaviour, while also monitoring teacher use of positive and negative statements. Similarly to Lannie and McCurdy, they found that positive statements did not increase during either the GBG or the CBGG, despite the CBGG's focus on positive behaviour. It should be noted however, that, in Wahl et al's evaluation of the CBGG, the teacher assigned and recorded points privately, only allowing students to know how many points they had received at the end of the game, similarly to the CBGG-d applied with adolescents in Study 1 in the current chapter. This could have impacted the teachers' rates of praise and negative statements. A comparison between teacher statements when the CBGG-d and CBGG-i are in place is an interesting and important line of inquiry. In Study 1 outlined earlier in this Chapter, there were minimal differences in student behaviour between the CBGG-d and the CBGG-i. Teacher statements were not manipulated, however it may be hypothesised that the teacher would engage in a higher rate of praise during the CBGG-i, due to the consistent visual feedback administered by the teacher. This could have important implications for understanding the mechanisms under which the game is effective.

The research questions for Study 2 remained the same as for Study 1, with a fifth research question added, as outlined next.

Research Questions

1. Is the CBGG an effective classroom management intervention in targeting academic engagement and disruptive behaviour in 'at-risk', mainstream populations of Irish adolescent students?
2. Is there a difference in effectiveness when the game is played with delayed feedback (CBGG-d) versus immediate feedback (CBGG-i)?

3. Do the implementing class teachers find the CBGG to be an acceptable intervention and do they prefer the CBGG-d or the CBGG-i?
4. Do students find the CBGG to be an acceptable intervention and do they prefer the CBGG-d or the CBGG-i?
5. Does the CBGG, in either format, impact teacher praise statements or reprimands?

Method

Recruitment, Participants and Setting

Participant recruitment was similar to Study 1 and took place in the same school but during the following school year. The first-year year head identified two first-year mathematics classes as potential target classes based on the high rates of teacher-reported disruption present and the importance of mathematics as a core subject in the Irish school system. The two mathematics teachers were informed of the study and given the opportunity to read about it in a PLS (Appendix M1). They then had the opportunity to ask questions during a short meeting and provided their consent to take part (Appendix M2). The teacher read an additional PLS and signed a separate consent form pertaining to the monitoring of teacher statements during the CBGG (Appendix M3; M4). Information packs were then arranged to be sent home with students in the relevant mathematics classes where parents were provided with PLS's and consent/assent forms for both parents and students (Appendix M5-M7).

Ms. Brady, was a 26-year-old female mathematics teacher with 3 years of experience, one of which was in the school in which data was collected. Her mathematics class consisted of 22 consenting students (10 female, 12 male) with a mean age of 12.7 years. Ms. Brady had not used contingencies like the CBGG before. Mr. Carroll was a 23-year-old male mathematics teacher, with 2 years of teaching experience. All of his experience was in the school in which data was collected. His mathematics class consisted of 16 consenting students (6 female, 7 male, 3 not reported), with a mean age of 12.8 years. Mr. Carroll had not used contingencies like the CBGG before but had seen the CBGG applied by another teacher (Ms. Allen) while team-teaching during the previous school year. Team-teaching in the school consisted of a main class teacher conducting primary class activities (Ms. Allen) while a 'team-teacher' (Mr. Carroll) provided additional support during the lesson.

Materials

Materials needed for the game were the same for both classrooms and were largely the same as, or similar to, those used in Study 1. Differences in materials used included both teachers changing from the smartphone app to a wearable smart watch reminder, with Ms. Brady using an Octopus watch (version 1) and Mr. Carroll using a Fitbit Charge 2. This change was made based on ease of use and more reliable prompts. The Octopus Watch is an icon-based watch designed to teach children good habits (Octopus Watch V2, 2020). It was chosen for use here because it could be pre-programmed to vibrate at regular intervals. The Fitbit Charge 2 is a fitness tracking device which can also be pre-programmed to vibrate at regular intervals (Fitbit Charge 2, 2020). The wearable devices were pre-programmed by the student researcher before the class began, meaning the teachers could leave their phones aside. Both devices were readily available to the student researcher in that they had been purchased for a range of other projects in the behaviour analysis lab in DCU. Prizes also differed slightly from Study 1. Prizes were identified using a preference assessment survey (similar to the survey used in Study 1; Appendix N) and highly rated prizes included homework passes, free time, sweets and school cinema passes. The daily and weekly leader boards were laminated A3 sheets, like those used in Study 1 (Appendix O) The checklist used by teachers in this study differed very slightly to that used in Study 1. Reviewing the game rules and explaining to students how to earn a point were divided into two steps whereas in Study 1 these were incorporated into one step. Therefore, there were 11 steps on the checklist used in this study (Appendix P). An additional data collection sheet was used by observers to collect data on teacher behaviour (Appendix Q), and the data sheets used to collect data on student behaviour remained the same as in Study 1.

Operational Definitions of Target Behaviours/Dependent Variables

Student Behaviour

Operational definitions for student behaviour remained identical to those used in Study 1. This was decided upon given the similarity of the class settings. Based on preliminary observations of the class they were deemed appropriate. Definitions of these behaviours are provided in the methods section for Study 1.

Teacher Behaviour

Teacher praise statements and reprimands were monitored but not targeted by any specific contingencies throughout the study. Teacher praise was operationally defined as

> *"Any teacher verbalisation directed towards a student/group of students positively acknowledging and praising rule following or responding to class content, for example, 'well done on raising your hand', 'good job, that's correct'".*

A teacher reprimand was defined as follows:

> *"Any teacher verbalisation directed towards a student/group of students expressing disapproval towards a student's behaviour or lack of rule following, including redirection of student behaviour, for example 'Put that pen down and concentrate on the board', 'Stop writing when you're supposed to be looking up at the board', 'stop swinging on your chair', 'shhhhh', 'put that drink away'".*

Observation Procedures and Data Collection

Data were collected for each class group up to five times per week during 40 min mathematics classes. Data collection for student behaviour followed identical methods to Study 1. Teacher behaviour was monitored via a frequency count in 30 s intervals across a 3-10 min period immediately before or immediately after data had been collected on student behaviour. Total frequency was then converted to rate per min. Data collection on student and teacher behaviour was counterbalanced such that data were collected first on students every second day. Data collection took place on consecutive school days, allowing for breaks due to school holidays/planned closures (e.g., Christmas holidays, mid-term break). There was one extended break in data collection in Ms Brady's class, where data collection could not resume after the Christmas break until the end of January (in the middle of the first withdrawal phase). This break was due to a student teacher taking over the class for three weeks in January. Data were considered carefully after this break and the data remained stable upon resumption. This was the only long break in data collection (i.e., >1 week) not related to a planned school closure.

Observer Training and IOA

Procedures for collection of IOA data on student behaviour were identical to those in Study 1 in terms of training, data collection and computations. Before the study commenced, observers were also trained in how to collect data on teacher behaviour (i.e., frequency data). IOA data were collected during 28.57% of sessions for Ms. Brady's class and 26.67% of sessions for Mr. Carroll's class. IOA for both student and teacher behaviour 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). Percentage IOA for teacher behaviour was calculated using the mean count per interval method. Percentage agreement was calculated for each 30 s interval, these percentages were added together and divided by the total number of intervals. If IOA fell below 80% for any observer, that observer was retrained before they were assigned to collect data again. Rates of IOA across both classrooms and all target behaviours are presented in Table 3.3. Mean IOA did not fall below 80% for any variable.

Table 3.3.

Mean IOA and Range Across Student and Teacher Behaviour

Class	AEB		DB		Teacher Praise		Teacher Reprimands	
	M	Range	*M*	Range	*M*	Range	*M*	Range
Ms. Brady's Class	86.59%	81.25-95.24%	90.65%	83.05-97.5%	99.44%	94.44-100%	90.21%	76.33-100%
Mr. Carroll's Class	80.89%	69.84-90%	80.10%	63.49-92.5%	93.98%	78.57-100%	81.96%	54.95-93.33%

Social Validity

Following the final day of data collection, the teachers and students completed social validity measures. The teachers completed the Behaviour Intervention Rating Profile (BIRS; Elliott & Von Brock Treuting, 1991; Appendix R) and students completed the same modified version of the CIRP used and described in Study 1 (CIRP; Mitchell et al., 2015; Witt & Elliott, 1985). 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 behaviour of the classroom). The BIRS incorporates the IRP-15 measure of acceptability used in Study 1

(Martens et al., 1985) and adds measures of efficiency and efficacy. Modifications were identical to those made in Study 1; items were modified to reflect application of an intervention to a group rather than an individual child, and to the present/past tense. Similar modifications were employed by other researchers evaluating the CBGG (e.g., Ford, 2017). Items were rated from 1 (strongly disagree) to 6 (strongly agree). Subscales on the measure assess a teacher's perception of intervention Acceptability (range 15-90), Effectiveness (range 7-42) and Efficiency (range 2-12). Higher scores indicate more positive perceptions.

Experimental Design

A withdrawal/reversal design with phases ABACABAC in Ms. Brady's class and ACABACABAB in Mr. Carroll's class was used to assess the effectiveness of the CBGG-d and CBGG-i. Phases were counterbalanced across the classrooms to control for potential order effects. Phase A refers to the baseline phase where regular teacher-planned contingencies were in place. Phase B refers to the CBGG-d and phase C refers to the CBGG-i, as in Study 1. This design was consistent with the WWC standards (WWC, 2017; 2020), incorporating at least three attempts to demonstrate intervention effects for both versions of the intervention. As in Study 1, phase lengths and changes were generally determined a-priori. If data were unstable toward the end of a phase, a suggestion would be brought to the teacher to extend the phase into the following week and a decision would be reached in conjunction with the teacher.

Procedure

Baseline

The baseline phase was identical to the baseline phase described in Study 1. Both teachers employed their usual disciplinary procedures while data collection was ongoing. Disciplinary procedures were similar to those used by Ms. Allen and included verbal warnings, sending students out of the room, penalty sheets, and sending students to their form teacher, year head or principal.

Teacher Training

Training took place for 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. As with Study 1, 5 min was decided as the most reasonable interval length. Classroom expectations were discussed with both teachers during training and both teachers decided to adopt the same set of expectations. These were: “I will present to class with all of my materials”, “I will raise my hand before speaking”, “I will respect my classmates and allow them to learn” and “I will follow my teacher’s instructions and remain on task”. The teachers and researcher discussed reasonable aims for points needed to obtain the prizes/reinforcers which would be available at the end of each game phase. These aims would be adjusted throughout intervention phases based on the students’ performance.

Intervention: Caught Being Good Game

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. For example, in Ms. Brady’s class, there were three distinct sections in the room where students sat, making three teams the most logical division. Apart from the number of teams, procedures for the game were identical across both classrooms and the same as in Study 1. The CBGG-d and the CBGG-i were introduced across different phases and were applied as described in Study 1.

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. 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 via email and they were encouraged to follow all intervention steps on the following days.

Mean treatment integrity was 79.55% (range 27.27-100%) for Ms. Brady and 72.73% for Mr. Carroll (range 27.27-100%). For Ms. Brady, the most commonly missed steps (those implemented less than 70% of the time) were generally towards the end of the game, and included reviewing the rules, announcing that the game had finished and reminding students how many points they needed to get the prize at the end of the game/announce the winners (if a Friday). These steps were implemented during between 56.25% and 68.75% of CBGG sessions. All other steps were completed 75% of the time or more. For Mr. Carroll, the most commonly missed steps were also generally towards the end of the game and included reviewing the rules, reminding students how many points they needed to get the prize, and the final three steps of the game (announce end of game, announce team points and write onto weekly scoreboard and remind students how many points they needed to get the prize at the end of the game/announce the winners). Mr. Carroll often filled in the points on the scoreboard at the beginning of the following class, however. All other steps were completed more than 72% of the time. As in Study 1, the steps most often missed were generally towards the end of the game session and the teachers tended to address these steps at the beginning of the following class, before commencing the next game session. For example, if the teacher missed the step where points were added to the overall weekly scoreboard at the end of a class session, this step was completed at the beginning of the following class period, before the next iteration of the CBGG. An overview of treatment integrity in each phase for each teacher is presented in Table 3.4.

Table 3.4.

Mean Rates and Range of Treatment Integrity during Intervention Phases for Ms. Brady & Mr. Carroll

Phase	Ms. Brady		Mr. Carroll	
	M	Range	*M*	Range
Intervention 1	81.82%	54.55-100%	87.88%	63.64-100%
Intervention 2	88.64%	72.73-100%	61.36%	27.27-81.82%
Intervention 3	70.45%	54.55-100%	74.03%	36.36-90.91%
Intervention 4	77.27%	27.27-100%	70.45%	54.55-81.82%
Intervention 5	-	-	72.73%	54.55-90.91%

Note. Phases are referred to as ‘Intervention 1’, ‘Intervention 2’ etc. in this table, as phase order for each classroom was different i.e., Intervention 1 refers to the CBGG-d for Ms. Brady and the CBGG-i for Mr. Carroll.

Data Analysis

As in Study 1, visual analysis of data was conducted in line with the WWC criteria for evaluation of single case research designs. Effect sizes were calculated using identical procedures to Study 1.

Ethical Considerations

Ethical considerations were the same as for Study 1 and additional considerations were undertaken for the monitoring of teacher behaviour (i.e., reprimands and praise statements) within the intervention study. A full new ethical approval form was prepared and submitted to DCU REC relating to the collection of data on teacher behaviour as it required assessment of additional risk. Only teachers who signed up and consented to take part in the larger CBGG intervention study would be considered to have their own behaviour monitored in the context of the CBGG. The teachers were not considered a particularly vulnerable population, however, there was a risk that they would find the monitoring of their own behaviour by a researcher uncomfortable. Before their behaviour was monitored, they read an additional PLS and completed an additional consent form. This documentation was separate from the CBGG intervention PLS and consent form and if they did not wish to consent to having their behaviour monitored, it would not affect their eligibility to take part in the CBGG study. The DCU REC approval letter for this ethical protocol is included in the Appendix (Appendix S).

Results

WWC Design Standards

As in Study 1, experiments in both classrooms in this study met the WWC design standards with reservations (WWC, 2017; 2020). The independent variable (i.e., the CBGG) was systematically manipulated. IOA data were collected at least once per phase, at least 20% of the time across phases and conditions and met the minimum thresholds of 80%. There were at least three attempts to demonstrate intervention effects for each version of the intervention. Each phase within the study contained at least 3-4 data points, meaning the study meets the standards with reservations. Two team-teachers were present in Mr. Carroll’s classroom towards the end of the study, which is a potential confound which will be discussed later in this section.

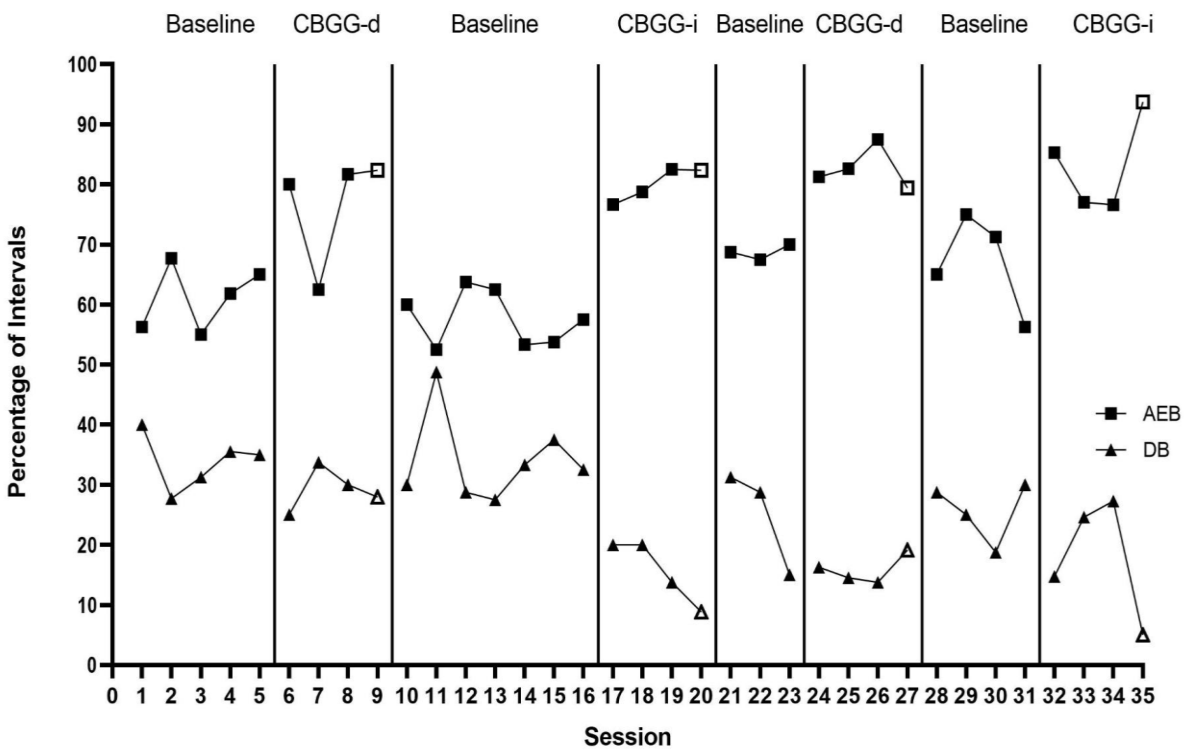
Student Behaviour

Ms. Brady's Mathematics Class

Figure 3.3. displays behavioural data for Ms Brady's class.

Figure 3.3.

Percentage of Intervals with AEB and DB in across Study Phases for Ms. Brady's Mathematics Class



Note. Open squares and triangles denote sessions where a weekly prize was available (i.e., a Friday)

Academically Engaged Behaviour. During the initial baseline phase, AEB was low in Ms. Brady's mathematics class, occurring during an average of 61.16% of intervals (range = 55-67.69%). When the CBGG-d was put in place, levels of AEB increased immediately and substantially, occurring during a mean of 76.63% of intervals (range = 62.5-82.35%). Although AEB dropped substantially during data point 7 (see Figure 3.3.), this may be attributable to low treatment integrity during this session (54.55%). Treatment integrity returned to 100% during the following session after the teacher received emailed feedback.

The overall increase in level of AEB upon introduction of the CBGG-d was moderate (an increase of 15.47%) and only data point 7 overlapped with the preceding baseline phase.

During the first withdrawal phase, AEB decreased immediately and substantially. It remained at a relatively stable and low level throughout this phase (*M* = 57.62%, range = 52.5-63.75%). The CBGG-i was put in place following the withdrawal phase and there was an immediate and large increase in AEB, which maintained a slight upward trend throughout the phase (*M* = 80.07%, range = 76.67-82.5%). There was no overlap between this phase and the previous withdrawal phase.

In the second withdrawal phase, AEB decreased immediately and remained stable across the phase (*M* = 68.75%, range = 67.5-70%). The CBGG-d was reinstated following this withdrawal phase. AEB increased immediately and remained high and stable across the phase (*M* = 82.69%, range= 79.41-87.5%). Again, no overlap was observed between this intervention phase and the withdrawal phase which preceded it.

During the final withdrawal phase, there was an immediate decrease in AEB when the intervention was withdrawn, however data was not as stable as in previous withdrawal phases. AEB occurred during a mean of 66.88% of intervals (range = 56.25-75%). Despite this, there was no overlap between this withdrawal phase and the preceding phase where the CBGG-d was in place. The CBGG-i was implemented in the final phase. There was an immediate and substantial increase in AEB (*M* = 83.18%, range = 76.62-93.75%) and there was no overlap with the preceding withdrawal phase.

Overall, rates of AEB were substantially higher in phases where either version of the CBGG was in place compared with baseline and withdrawal phases. Only one data point (data point 7; Figure 3.3.) in an intervention phase overlapped with data points in the baseline/withdrawal phases. Mean rates of AEB did not exceed 75% in baseline and withdrawal phases, whereas mean rates in intervention phases always exceeded 76% and mean rates in three of four intervention phases exceeded 80%.

Disruptive Behaviour. DB was high during the initial baseline phase, occurring during a mean of 33.89% of intervals (range = 27.69-40%). 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* = 29.17%, range = 25-33.75%). Three of the four data points overlapped with baseline.

When the CBGG-d was withdrawn, changes in DB were not very pronounced and remained at a level similar to the previous phase (*M* = 34.05%, range= 27.5-48.75%). Upon introduction of the CBGG-i for the first time, there was an immediate and substantial decrease in DB (*M* = 15.64%, range = 8.82-20%), 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 across two data points, however decreased again towards the end of the phase (*M* = 25%, range = 15-31.25%). When the CBGG-d was reinstated, DB remained low and stable across the phase (*M* = 15.9%, range = 13.75-19.12%), however 50% of the data points overlapped with the preceding withdrawal phase.

DB increased modestly to a mean of 25.63% of intervals (range = 18.75-30%) during the following, final withdrawal phase. The CBGG-i was implemented during the final phase and DB decreased immediately, occurring during a mean of 17.89% of intervals across the phase (range = 5-27.27%). However, across the phase, two data points saw DB increase to levels similar to the previous withdrawal phase, which resulted in a 50% overlap.

Overall, there were some improvements evident in DB when the CBGG was in place versus baseline/withdrawal phases. Mean rates of DB always decreased in CBGG phases compared to the preceding baseline/withdrawal phases, with the largest mean decrease occurring between the first withdrawal phase and the introduction of the CBGG-i for the first time (18.41% decrease). Changes in DB were not as stable or substantial as changes in AEB in this class group.

Effect Size Data. Tau effect sizes across phase changes in Ms. Brady's class are presented in Table 3.5. 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 and -.64 respectively. The effect size was therefore moderate for the CBGG-d and large for the CBGG-i. Overall, the CBGG-i produced larger effect sizes for both target behaviours across phase changes.

Table 3.5.

Tau Effect Sizes Across Phase Changes for Ms. Brady's Mathematics Class

Phase Change	AEB	DB
Baseline to CBGG-d 1	.596	-.447
Withdrawal to CBGG-i 1	.714*	-.72*
Withdrawal 2 to CBGG-d 2	.756	-.504
Withdrawal 3 to CBGG-i 2	.756*	-.472
Weighted Mean CBGG-d	.68	-.47
Weighted Mean CBGG-i	.73	-.64

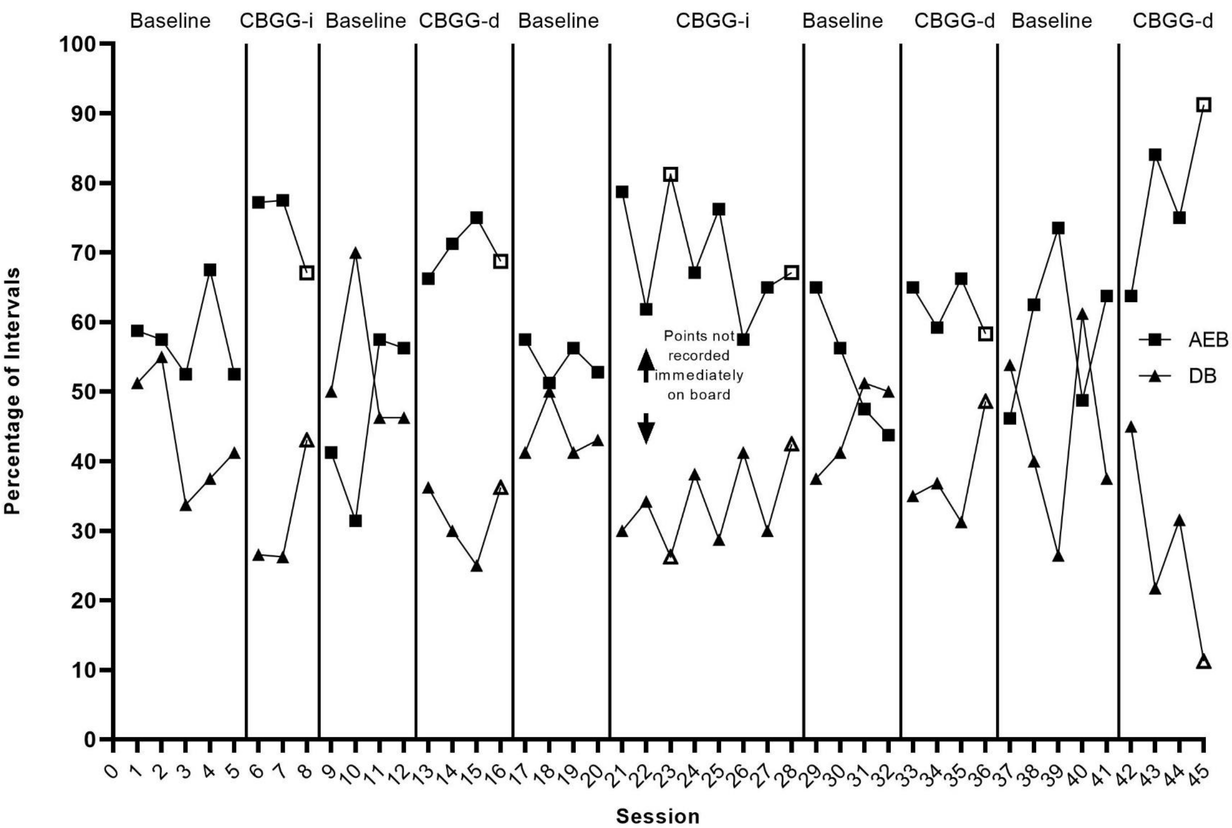
*$p<.05$

Mr. Carroll's Mathematics Class

Figure 3.4. is a graph depicting AEB and DB across phases in Mr. Carroll's mathematics class.

Figure 3.4.

Percentage of Intervals with AEB and DB Across Study Phases for Mr. Carroll's Mathematics Class



Note. Open squares and triangles denote sessions where a weekly prize was available (i.e., a Friday). During data point 22, Mr. Carroll recorded points privately in error rather than publicly.

Academically Engaged Behaviour. AEB in Mr. Carroll's class was relatively stable at baseline, occurring during a mean of 57.75% of intervals (range = 52.5-67.5%). When the CBGG-i was implemented for the first time, there was an immediate increase in AEB (*M*= 73.94%, range = 67.09-77.5%). 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 was the only one to overlap with the baseline phase.

Despite a decrease in AEB towards the end of the previous intervention phase, there was an even larger decrease in AEB at the beginning of the first withdrawal phase (*M* = 46.61%, range = 31.43-57.5%). The CBGG-d was then put in place for the first time and there was an increase in AEB (*M* = 70.31%, range = 66.25-75%) 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* = 54.45%, range = 51.25-57.5%) 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 3.3.). To ensure enough data were collected on the CBGG-i during the phase (i.e., >3 points), the phase was extended across two weeks. There was a change in level across AEB (*M* = 69.35%, range = 57.5-81.25%), however data were highly variable across this phase, with a general downward trend for AEB. Despite this undesirable trend and highly variable data, only one data point (data point 26; see Figure 3.4.) overlapped with the previous phase, meaning there was a high percentage of non-overlapping data (87.5%).

During the next withdrawal phase, AEB continued to decrease at a much steeper rate than during the CBGG-i (*M* = 53.13%, range = 43.75-65%). There was an immediate improvement evident in AEB when the CBGG-d was put in place for the second time (*M* = 62.2%, range = 58.33-66.25%) 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 extra baseline phase was put in place, before trialling the CBGG-d a final time with a different prize.

Behaviour was highly variable during the final withdrawal phase. AEB occurred during a mean of 58.94% of intervals (range = 46.15-73.53%). During the third data point in this phase (point 39; Figure 3.4.), there were two team teachers present alongside the main class teacher. 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* = 78.52%, range = 63.75-91.25%). However, changes were not immediate and there was a degree of overlap with the previous phase. During this week, students were approaching a long weekend break and absenteeism was high. As well as this, two team teachers were present again during data points 43 and 45 (Figure 3.4.). Therefore only tentative conclusions must be drawn for this phase change due to the confounding variables at play.

Disruptive Behaviour. During the initial baseline phase, DB was variable, however was demonstrating an upwards trend towards the end of the phase, across the final three data points (*M* = 43.75%, range = 33.75-55%). The CBGG-i was introduced and DB decreased to a mean of 31.96% of intervals (range = 26.25-43.04%). The data did not remain stable however, and the final data point in this phase saw an increase in DB which led to overlap with the baseline phase.

The CBGG-i was withdrawn and there was an increase in DB (*M* = 53.13%, range = 46.25-70%). There was no overlap here with the preceding intervention phase. The CBGG-d was introduced during the following phase. DB decreased immediately (*M* = 31.88%, range = 25-36.25%). DB remained moderately stable and there was no 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* = 43.89%, range = 41.25-50%). 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 3.4.). DB was low initially in this phase, however increased across the phase and remained unstable (*M* = 33.89%, range = 26.25-42.47%). Data points 26 and 28 overlap with the preceding withdrawal phase, meaning the percentage of non-overlapping data is 75%.

Although DB did not increase immediately during the following withdrawal phase, behaviour increased steadily across the phase (*M* = 45%, range = 37.5-51.25%). The CBGG-d was then put in place and an immediate decrease in DB was evident (*M* = 37.93%, range = 31.25-48.61%). 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 43.81% of intervals (range = 26.47-61.25%) during the final withdrawal phase. The data were very 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 during the final week (*M* = 27.39%, range = 11.25-45%), however this must be considered in light of the confounds reported in the AEB section above.

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 3.6. When calculating these effect sizes, 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.

Table 3.6.

Tau Effect Sizes Across Phase Changes for Mr. Carroll's Mathematics Class

Phase Change	AEB	DB
Baseline to CBGG-i 1	.646	-.439
Withdrawal to CBGG-d 1	.756*	-.784*
Withdrawal 2 to CBGG-i 2	.701*	-.582*
Withdrawal 3 to CBGG-d 2	.529	-.567
Withdrawal 4 to CBGG-d 3	.643*	-.447
Weighted Mean CBGG-d	.66	-.64
Weighted Mean CBGG-i	.68	-.53

**p<.05*

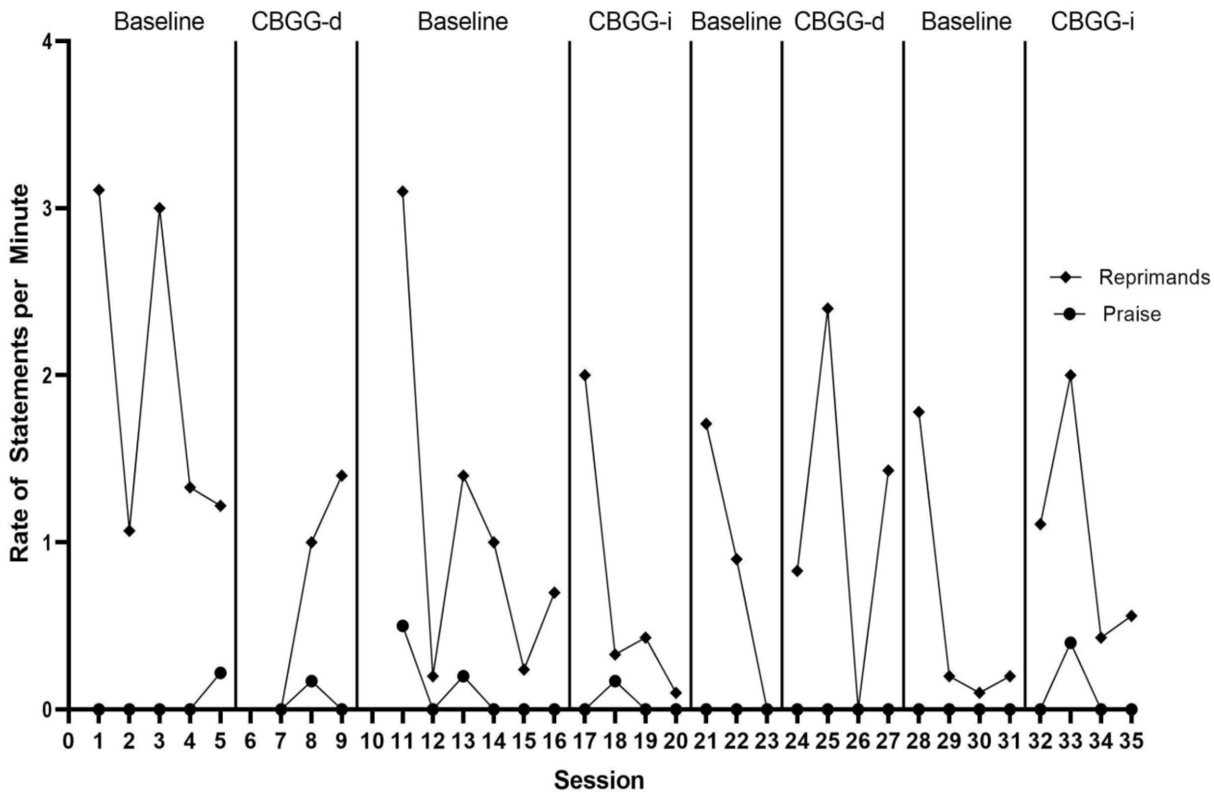
Teacher Behaviour

Ms. Brady

Figure 3.5. displays teacher behaviour data for Ms. Brady.

Figure 3.5.

Ms. Brady's Rate of Praise Statements and Reprimands per Minute Across Study Phases



Praise Statements. Ms. Brady's rate of praise statements was low and stable rate across all study phases. During the first baseline phase, praise statements occurred at a mean rate of 0.04 statements per min (range = 0-0.22). No change was evident when the CBGG-d was introduced, with praise statements occurring at a mean rate of 0.06 times per min (range = 0-0.17). Although praise occurred at a mean rate of 0.12 statements per min during the first withdrawal phase (range = 0-0.5), this increase in the mean rate is attributable to one data point where praise occurred at a rate of 0.5 times per min. There were four data points in this phase where no praise statements were used by the teacher during the observation session. The introduction of the CBGG-i did not lead to any increases in praise rates. Praise statements occurred at a mean rate of 0.04 times per min (range = 0-0.17). In the following three phases (baseline, CBGG-d, baseline), no praise statements were recorded during the observation sessions. In the final phase (CBGG-i), praise occurred at a mean rate of 0.1 times per min (range = 0-0.4). Overall, praise occurred at a very low rate across all phases and the experimental condition did not impact the rates observed.

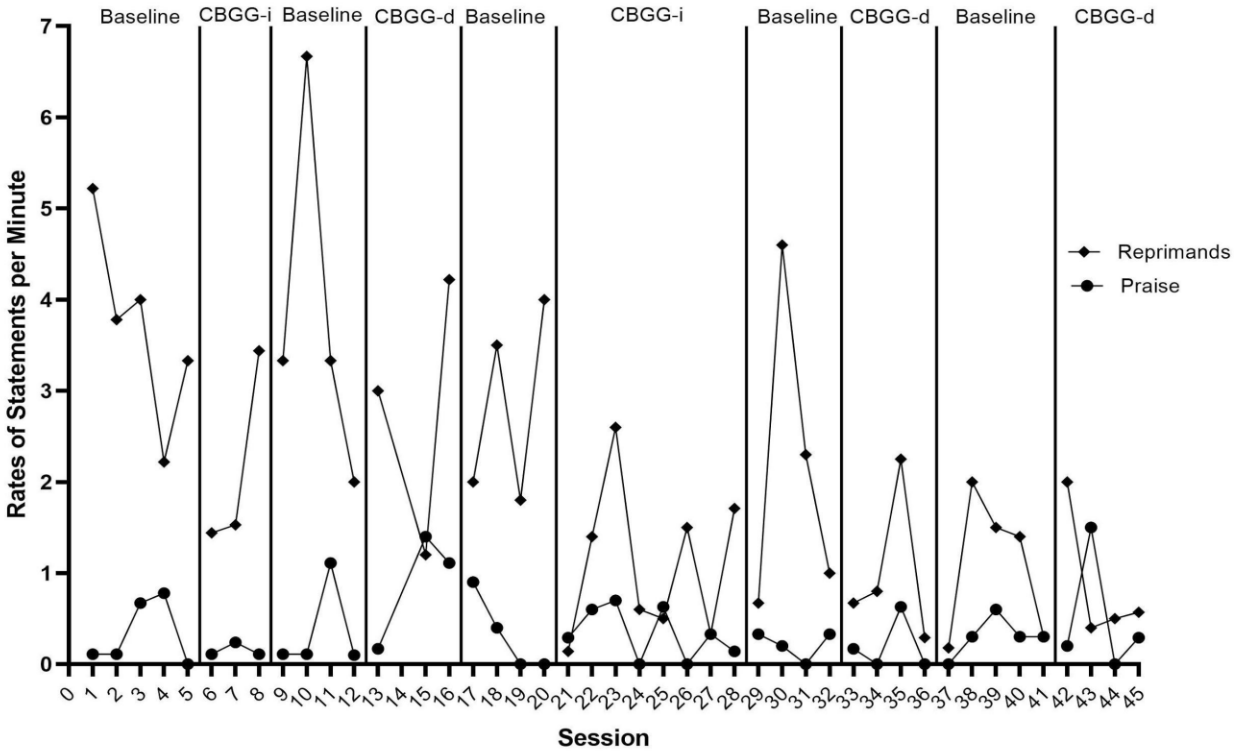
Reprimands. At baseline, reprimands occurred at a mean rate of 1.95 statements per min (range = 1.07-3.11). Although there was an apparent decrease when the CBGG-d was introduced (*M* = 0.8, range = 0-1.4), statements increased to a level similar to baseline by the end of the phase. During the first withdrawal phase, although there was a sharp increase in the beginning, the mean rate of teacher reprimands was 1.11 statements per min (range = 0.2-3.1). Introduction of the CBGG-i was accompanied by another increase in reprimands compared to the final data point in the previous phase, however this stabilised at a lower rate throughout the remainder of the phase (*M* = 0.72, range = 0.1-2). Rate of teacher reprimands increased slightly to 0.87 statements per min during the next withdrawal phase, with a decreasing trend apparent across the phase (range = 0-1.71). The second iteration of the CBGG-d saw a slight increase in reprimands to 1.17 statements per min (range = 0-2.4). During the final withdrawal phase, teacher reprimands decreased to a rate of 0.57 statements per min (range = 0.1-1.78). During the final iteration of the CBGG-i, reprimands occurred at a mean rate of 1.03 statements per min (range = 0.43-2). In sum, reprimands were largely variable across all study phases, with no discernible pattern attributable to intervention implementation.

Mr. Carroll

Figure 3.6. displays teacher behaviour data for Mr. Carroll.

Figure 3.6.

Mr. Carroll's Rate of Praise Statements and Reprimands per Minute Across Study Phases



Praise Statements. Praise statements occurred at a low rate across all experimental phases. During the initial baseline phase, praise statements occurred during a mean rate of 0.33 times per min (range = 0-0.78). When the CBGG-i was put in place, the mean rate decreased to 0.15 statements per min (range = 0.11-0.24), but this decrease was not substantial. During the first withdrawal phase, the mean rate of praise statements remained low (*M* = 0.36, range = 0.1-1.11). There was a slight increase in praise statements when the CBGG-d was put in place (*M* = 0.89, range= 0.17-1.4), however, this increase was minimal. Mean rate of praise statements decreased steadily across the following baseline phase until it occurred at a rate of 0 during the last two data points (*M* = 0.33, range= 0-0.9). Rate of praise statements was stable and low during the following lengthy CBGG-i phase (*M* = 0.34, range = 0-0.7). This pattern of little variance was maintained across all following study phases; baseline (*M* = 0.22, range= 0-0.33), CBGG-d (*M* = 0.2, range= 0-0.63), baseline (*M* = 0.3, range = 0-0.6) and finally CBGG-d (*M* = 0.5, range = 0-1.5). The implementation of the CBGG in either format did not appear to impact Mr. Carroll's praise rates.

Reprimands. Mr. Carroll used reprimands at a mean rate of 3.71 statements per min at baseline (range = 2.22-5.22). A decreasing trend was evident across this phase. Although there was a slight decrease when the CBGG-i was introduced, an increase was evident again during the last data point in the phase (*M* = 2.14, range = 1.44-3.44), corresponding with an increase in DB during this observation (data point 8; Figure 3.4.). During the first withdrawal phase, rate of teacher reprimands per min was highly variable, occurring during a mean rate of 3.83 per min (range = 2-6.67). Variability was also present in the next CBGG-d phase. Teacher reprimands occurred at a mean rate of 2.81 per min (range = 1.2-4.22). During the second withdrawal phase, teacher reprimands occurred at a similar rate to the previous phase (*M* = 2.83, range = 1.8-4). A decrease in overall level was evident when the CBGG-i was put in place for a second time. Reprimands occurred at a mean rate of 1.1 statements per min (range = 0.14-2.6). In the following baseline phase, there was a slight increase with reprimands occurring at a mean rate of 2.14 times per min (range = 0.67-4.6). The CBGG-d was put in place for a second time and reprimands decreased slightly to a mean rate of 1 per min (range = 0.29-2.25). During the final withdrawal phase the mean rate remained similar (*M* = 1.08, range = 0.18-2). The CBGG-d was put in place during the final experimental phase, and there was a very slight decrease in mean rate of reprimands (*M* = 0.87, range = 0.4- 2). In sum, rates of teacher reprimands were very variable across all study phases. There were some decreases evident, particularly when the CBGG-i was in place, however any impact was small.

Social Validity

Teacher Social Validity

Teacher feedback on the BIRS (Elliott & Von Brock Treuting, 1991) was conflicting. Ms. Brady responded predominantly negatively, rating the intervention 37 on Acceptability (*M* = 2.47), 7 on Effectiveness (*M* = 1) and 2 on Efficiency (*M* = 1). 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 behaviour 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 77 for Acceptability (M = 5.13), 28 for Effectiveness (M = 4) and 10 for Efficiency (M = 5). In his written feedback he stated that he thought the game was "*definitely a worthwhile intervention*" and that "*the students did react positively*". He also 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 completed the modified CIRP (Mitchell et al., 2015; Witt & Elliott, 1985). They rated the CBGG moderately positively on the CIRP (M = 4.79, range = 1-8). A majority (i.e., 10 or more) students responded positively to at least six of the eight statements. Twelve students disagreed that the game helped them to do better in mathematics class and ten students thought the game may have caused problems for their classmates. Few students wrote additional comments on the game, and no student elaborated on the problems that they perceived it may have caused for classmates. Problems likely related to teams sometimes not winning a prize, or individual students sabotaging a team's chances at winning a prize. 42.11% of students preferred the CBGG-i, 42.11% preferred the CBGG-d and two students did not report a preference.

Fourteen students in Mr. Carroll's class completed the modified CIRP (Mitchell et al., 2015; Witt & Elliott, 1985). They rated the game more positively than Ms. Brady's class, scoring a mean of 6.36 (range= 3-8). 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 which could have been caused by the game. In denoting a preference, 64.29% of students preferred the CBGG-i and 28.57% preferred the CBGG-d. One student did not report a preference.

Summary & Brief Discussion Study 2

The current study aimed to build upon Study 1 by implementing the CBGG-d and the CBGG-i across two first-year class groups, while counterbalancing the presentation of intervention conditions. It sought to further probe the effectiveness of the CBGG-d and the CBGG-i with first-year mathematics students and to assess teacher statements during the intervention. Social validity was also an important element of this study, and ratings were obtained from part-taking teachers and students.

Effectiveness of the CBGG with and without Visual Feedback

The CBGG was generally successful in targeting AEB and DB across the two classrooms in this study. In Ms. Brady's class, AEB was almost always higher in phases where the CBGG was in place compared to 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 in Ms. Brady's class (.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 behaviour when compared to baseline, and this is reflected in a moderate and non-significant effect size for the phase change (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. Changes in DB were not as stable or substantial in this class group when compared to changes in AEB. This is reflected in slightly smaller effect sizes for the CBGG-d (Weighted Mean Tau = -.47) and the CBGG-i (Weighted Mean 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; Figure 3.4.), 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 Tau = .66) and CBGG-i (Weighted Mean Tau = .68) on AEB and differences between the two versions of the game were minimal. Effects of the CBGG on DB were mainly positive, with some mixed results. There was always a decrease in mean DB when a game phase was introduced versus the preceding baseline or withdrawal phase. Towards the end of the study, behaviour 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 programme 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 towards the end of the study, which may have influenced student DB.

Taken together, the above results align with the findings of Study 1 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 behaviour (e.g., Ford, 2017; Tanol et al., 2010; Wahl et al., 2016; Wright & McCurdy, 2012). The current results support findings from Study 1 in the current chapter in demonstrating the efficacy of the game across two additional first-year classes. Feedback was again manipulated, and the game was effective in both formats, lending support to the efficacy of the game with delayed feedback (e.g., Wahl et al., 2016) and immediate visual feedback (e.g., Lynne et al., 2017). The importance of this finding will be outlined in the general discussion section later in this Chapter. The study design was significantly stronger here compared to Study 1 and allowing for counterbalancing of the intervention conditions means that order effects could be controlled for. This type of scientific rigour is important in single-case research so that confident deductions can be made about the effects of the intervention on behaviour. Study 1 alone was not enough to draw strong conclusions and Study 2 was developed based on this limitation.

Social Validity of the CBGG with Teachers and Students

Ms. Brady rated the game predominantly negatively, citing time constraints as a critical issue when giving reasons for her unfavourable rating. Mr. Carroll rated the game very positively and called it a "*worthwhile intervention*" for his class. 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 (ASTI, 2018). It is possible that Ms. Brady was experiencing this and that engaging in research on a classroom management intervention was, in her words, adding an "*extra layer of work"*. 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., 2019) would be more desirable for Ms. Brady. Mr. Carroll rated the CBGG positively in terms of acceptability, effectiveness, and efficiency. This positive rating aligns with the rating given by Ms. Allen in Study 1. Both Ms. Brady and Mr. Carroll preferred the CBGG-d to the CBGG-i, which is in agreement with Ms. Allen in Study 1. The key reasoning behind these preferences appears to be the distraction element in that the CBGG-d is less distracting for teachers 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 teacher ratings, in that Mr. Carroll rated the game more positively than Ms. Brady. The positive ratings align with much previous research in this area, in that students generally tend to rate the GBG and CBGG favourably (e.g., Kleinman & Saigh, 2011; Wahl et al., 2016; Wright & McCurdy, 2012). Mr. Carroll's class rated the game very similarly to Ms. Allen's class.

Teacher Praise and Reprimands during the CBGG

Teacher praise rates were low for both teachers across the entire study and did not change depending on the study phase. Rates of praise were invariable in that they remained very low across all phases, usually not exceeding a rate of 1 per min. This finding is in line with previous research which found that implementation of the CBGG did not impact teacher praise levels, and that teacher negative statements tended to be higher than praise statements (Wahl et al., 2016). This finding may indicate a need to incorporate a teacher contingency to encourage use of teacher praise throughout CBGG implementation, given its research-based efficacy in having a positive influence on student behaviour (Moore et al., 2019). Data on teacher behaviour were only collected for up to 10 min during a class period, so this must be taken into account when evaluating the results. It is possible that there were no praise statements during the data collection session because quiet, independent study was the ongoing activity.

Rates of teacher reprimands did not appear to be strongly impacted by the implementation of either the CBGG-d or the CBGG-i. Although there were instances where reprimands decreased when an intervention phase was instigated, data generally remained highly variable and there were large rates of overlap. Ms. Brady's rate of reprimands varied greatly from phase to phase, with high rates present during both baseline/withdrawal and intervention phases at times. Mr. Carroll's rate of reprimands were slightly higher than Ms. Brady's overall, however this is not surprising given his class group were less engaged and more disruptive than Ms. Brady's class in general. It must be reiterated that data on teacher behaviour were only collected for up to 10 min during data collection sessions and this therefore limits conclusions which can be drawn from the data. As student behaviour was the main focus during this study, teacher behaviour was not assigned as much time.

General Discussion

The studies outlined in the current chapter aimed to assess the efficacy of the CBGG both with and without immediate visual feedback across three populations of mainstream Irish secondary school students. Study 1 compared the CBGG-d with the CBGG-i across one first-year mathematics class, using an ABACABAC reversal/withdrawal design. Study 2 built upon Study 1, by comparing the CBGG-d and CBGG-i across two further first-year mathematics classes, counterbalancing intervention conditions. Study 2 also examined wither either version of the intervention impacted teacher rates of praise or reprimands. Taken together, the results of the studies provide promising evidence that the CBGG is an effective intervention with this population, and that feedback is not a particularly crucial component. It is important to note however, that social validity data indicated that students tend to prefer the game when feedback is given immediately, whereas teacher prefer implementing the game with delayed feedback.

The Effectiveness of the CBGG-d and the CBGG-i with Adolescent Students

The current findings are important in many respects. With reference to research questions one and two, the findings demonstrate that the CBGG is potentially useful with an adolescent student population in an Irish context and that the CBGG-d was as effective as, and sometimes more effective than the CBGG-i. This is a particularly important finding for teachers who are under time pressure and may not have flexibility in their schedule to engage in intensive behaviour support plans or individualised classroom management practices. Few published studies to date have attempted to determine if feedback was a necessary component of the traditional GBG and none have evaluated feedback as a flexible component during the CBGG. Harris and Sherman (1973) concluded that feedback was not a necessary component during GBG implementation. The current results provide further evidence for this finding 47 years after that study by Harris and Sherman, and with an updated, positive version of the game. Wiskow et al., (2019) found that feedback was important during the GBG and that visual + vocal feedback produced the largest changes in behaviour with a preschool class. The GBG with no feedback was not effective until after the game had been put in place with feedback. These results must be interpreted in context of the population recruited however; preschool children may not have requisite skills to understand visual feedback in that they may be unable to count marks. Therefore, exploration of feedback with adolescent students was an important line of inquiry, given their level of understanding compared to preschool populations. Foley et al., (2019) also examined the feedback component of the GBG, however did so as part of a larger component analysis and therefore did not isolate feedback as a changeable component itself. This study was also conducted with preschool students. The current studies address several limitations of the above studies by examining the component of feedback in the context of the CBGG (rather than the GBG) and by evaluating feedback with adolescent populations.

Social Validity of the CBGG

Regarding research questions three and four, teacher and student ratings of the CBGG differed across the three classrooms. Ms. Allen and Mr. Carroll provided similar ratings and rated the game positively, whereas Ms. Brady rated the game much more negatively. Ms. Brady’s negative appraisal appeared to relate to time constraints, many of which were outside the scope of control for this research. The students in Ms. Allen’s class and Mr. Carroll’s class rated the game similarly and preferred the CBGG-i, whereas Ms. Brady’s class did not prefer one version over the other. The general student preference for the CBGG-i is not surprising. Positive reinforcement in the form of points likely formed part of the participants’ learning history. The students had likely taken part in visual points-based reinforcement systems in other classes or encountered positive reinforcement by earning points during sports or other game participation. For this reason, students rating the CBGG-i more positively may have felt more comfortable with the familiarity of knowing how many points they had at a given time. These findings make it difficult to assert which version of the game should be applied with a similar group in future as the part-taking teachers all preferred the CBGG-d. It may be useful to examine whether teachers would be willing to implement the game first with feedback and then fade feedback over time to save on resources.

Teacher Praise during the CBGG

Teacher statements were evaluated in Study 2 and those findings align with previous research which found that the GBG and CBGG do not greatly impact teacher positive statements (Lannie & McCurdy, 2007; Wahl et al., 2016). Lannie and McCurdy (2007) previously found that when DB decreased during the GBG, teacher neutral and negative statements concomitantly decreased, however the same was not observed in Study 2. Wahl et al., (2016) found that no increases in positive statements occurred during the CBGG or GBG despite hypothesising that the CBGG may lead to increases due to its focus on positive behaviour. Immediate feedback was not given to students during the CBGG in the study by Wahl et al., and this may have meant the teacher did not want to draw attention to behaviour, either positive or negative, while awarding points. The current study evaluated both delayed and immediate feedback and has demonstrated that even with immediate feedback, teachers did not engage in a higher rate of praise.

The potential impact for praise during interventions like the CBGG should not be understated. Praise has positive effects on appropriate and disruptive student behaviour when applied as an intervention (Moore et al., 2019), therefore encouraging high levels of praise during interventions like the CBGG, could potentially elevate the intervention effects. Despite positive connotations of praise, there have been mixed reactions among students as to their preferences around praise. Elwell and Tiberio (1994) found that students generally reported that they thought they should be praised for behaving well in class, however when asked specifically if they would like to be praised for being well behaved, 54% stated they would prefer no praise at all to loud praise (9%) or quiet praise (37%). In a more recent study, Fefer et al, (2016), found that 82.6% of adolescent students thought they should be praised at least sometimes for doing their work, with most students preferring quiet praise to public/loud praise. Despite adolescents' mixed attitudes towards praise, Blaze et al., (2014) demonstrated that systematically manipulating teacher rates of loud and quiet praise across different study phases have similar, positive outcomes on academic engagement and disruptive behaviour among students in ninth to eleventh grades. Four teachers took part in this study by Blaze et al.; two preferred loud praise and the other two preferred quiet praise. These findings regarding praise preferences point towards the need for consideration of the teacher and students under study before manipulating praise during the CBGG. Evaluation of the game with deliberate loud praise versus deliberate quiet praise may be an avenue worth exploring in future research. Strategies such as prompting (Blaze et al., 2014) or self-monitoring may be useful in encouraging teacher to incorporate more praise during the CBGG.

Conducting Research in the Classroom Environment

Across the three classrooms, teacher treatment integrity was low on occasion. Potential reasons for low treatment integrity have been identified in previous research. Wehby et al., (2012) have suggested that teacher-coach working alliance and teacher's perception of social validity can impact upon treatment fidelity during GBG implementation. The researcher in these studies did not serve as a 'coach' as such but did provide the teacher with support and contact throughout the process. Teachers over the age of thirty have also been found to implement the GBG with less fidelity than teachers under thirty (Domitrovich et al., 2015). The same authors found that teachers were less likely to choose to implement the GBG at all if it did not align with their teaching style. Reduction of general workload stress has also been found to coincide with higher levels of evidence-based intervention implementation fidelity (Larson et al., 2018). These factors should be taken into account in future iterations of the game.

Another issue with the naturalistic setting of a classroom is that there were sometimes days where data collection was due to take place and cancelled at short notice. In the current studies this was due to incidences like parent-teacher meetings and severe weather conditions leading to school closures. This sometimes led to a break in gameplay for several days. This was potentially problematic as prizes were to be given out weekly. If a data collection session was missed inadvertently, the researcher and teacher would have to discuss altering the points goal and the teacher would then inform the students of any alterations. This could have caused some confusion among students, however the teachers always strived to be clear about any alterations to the procedures Unplanned breaks in data collection sessions also meant that sometimes where it was planned that a phase would contain the five data points recommended by the WWC (2017; 2020), only three or four were collected.

Weekly Prizes During the CBGG

An interesting ancillary finding of the current research was the apparent efficacy of the CBGG with prizes only awarded weekly. This variation was introduced based on teacher concerns with awarding prizes daily after class when time was already short to get class work completed. Previous iterations of the GBG and CBGG with secondary school students have applied the game for a certain period of a class (e.g., 20 min; Dadakhodjaeva, 2017), or throughout a whole class period with prizes awarded at the end (Ford, 2017). Kleinman and Saigh (2011) applied the GBG with a secondary school history class and found that awarding small daily prizes (sweets) and a large weekly prize (pizza party) was effective. The procedural variation employed in these studies (weekly prizes only) still produced positive intervention effects across three first-year classrooms, which has further important implications for teachers who may be under time pressure in secondary school settings. This procedural variation paired with delayed visual feedback may reduce the workload during the CBGG substantially when compared with procedures used in previous studies. Future research may consider comparing a version of the CBGG where prizes are awarded daily with a version where prizes are awarded weekly to assert whether the weekly prizes are more effective, as effective or less effective than daily prizes with this cohort.

Limitations

Study 2 attempted to address many previously discussed limitations for Study 1 by recruiting two classes and counterbalancing intervention conditions. However, in considering the two studies, there are some remaining limitations of note. First, results are relevant to first-year mathematics class settings. No other year-groups or classes part-taking in other subjects were recruited during this study which limits the generalisability of the results. As well as this, across the two studies, data were only collected on the classes as a whole. It was evident in teacher evaluations that sometimes individual students may not have engaged fully with the CBGG and may have needed more intense behaviour strategies. Third, only visual feedback was evaluated across these studies. Although rates of teacher statements were monitored, vocal feedback was not manipulated or examined here as it was in a previous study on feedback during the GBG by Wiskow et al., (2019). A final limitation relates to the materials used during this piece of research. The Octopus watch used by Ms. Brady is designed for children and is therefore small and may not be suited for use by some teachers. It was used based on convenience for the study as it was available as a resource within the research lab. Therefore, although it served the purpose for which it was intended in this study, future research may consider use of alternative devices. The Fitbit Charge 2 (and its successors/updated versions) may be a more useful device which teachers have access to, given the popularity of smart wearables.

Implications for Future Research and Practice

A major implication of this study is that the CBGG in both formats was successful in the decrease of DB and increase of AEB across first-year class groups. The studies serve as a platform for future research on the use of the CBGG with Irish secondary school students. It has evidently been effective over time with three first year class groups, which suggests that it may be successful with other first-year class groups and perhaps older class groups in a secondary school setting. Although the CBGG had been tested previously with secondary school students, the delayed feedback version had not been, and previous iterations involved the incorporation of Class Dojo technology (Dadakhodjaeva, 2017; Ford, 2017). This simpler version may be more desirable for use by teachers in secondary school classrooms who have a limited time frame in which to get work done. The teachers in these studies preferred the CBGG-d, so it is likely that they would have preferred this version over a version incorporating Class Dojo.

Conclusion

Either version of the CBGG applied in this study may be considered by teachers in lower secondary school settings. This is particularly relevant as schools adopt more positive behavioural approaches, as the CBGG maintains a focus on encouraging rule following rather than punishing or reacting to rule breaking. Comparison of two versions of the game across these two studies was an important line of inquiry which adds to the growing body of literature on procedural variations of the CBGG. Of relevance here is that the variation in question (i.e., the withholding of feedback) may lead to the CBGG being more adaptable by teachers given the time it may save for them in implementation. It is likely that teachers may prefer using the CBGG-d as it saves some time, however further research is needed on its effectiveness when compared with a version of the game which is more similar to the classic GBG. It is suggested that future research focuses on replicating application of the CBGG with Irish students, examining variable, potentially time-saving components and focusing on how the intervention impacts individual as well as group behaviour.

Chapter 4: The Caught Being Good Game in a Single-sex Primary School Setting

Introduction

The effectiveness of the GBG and CBGG has been well-established in mainstream school settings, however studies have typically focused on mixed-sex classes of students attending school in the USA (e.g., Dadakhodjaeva et al., 2019; Donaldson et al., 2018; Kleinman & Saigh, 2011; Lynne et al., 2017; Mitchell et al., 2015; Wright & McCurdy, 2012). The generalisability of these findings to novel populations needs to be tested. Chapter 3 was an important first step in addressing this through its focusing on the use of the CBGG with an understudied adolescent population however, there is still considerable scope to examine the use of the CBGG further. One area of particular interest is to examine the effects of a class based CBGG on the behaviour of individual students. In addition, and perhaps of importance in countries like Ireland, where single-sex education is common, is investigating the use of the CBGG in a single-sex mainstream school environment. These two issues are the focus of the study reported in the current chapter.

Evaluating Game-based Interventions with Groups vs Individuals

During observations of class-wide classroom management interventions such as the CBGG, it is common that researchers only collect and analyse data on the group as a whole (e.g., Wahl et al., 2016; Wright & McCurdy, 2012). Whole class data may ‘mask’ non-responders, which is an issue when it comes to evaluating the true effects of an intervention. It is important to understand whether general, class-wide behavioural interventions are being effective with individual students, particularly those perceived by a class teacher to be more disruptive than others. If a whole class responds well to an intervention like the CBGG but one individual student who usually engages in a high rate of disruptive behaviour does not respond well, this can lead to several potential difficulties. First, to the observers, the game appears effective across the whole class, however it may have resulted in behaviour change for the majority of the students while having no effect for an important minority. Second, the game affects behaviour change for most students in the class but is perceived as ineffective by the teacher due to one or two students still engaging in a high rate of disruption. This could in turn lead to low social validity ratings from the teacher despite some improvements being evident. By collecting data on individuals within the context of the whole class group, there is the potential to identify students who need more intensive behavioural interventions than the one being implemented (Donaldson et al., 2017). This may be particularly relevant in school settings where students are deemed ‘at-risk’, such as DEIS schools in an Irish context.

In the review presented in Chapter 2, three studies were identified which evaluated the effects of a game-based intervention on individuals only, while five studies evaluated both individual and group behaviour. Of these eight studies, two evaluated games similar to the CBGG (Conklin et al., 2017; Tanol et al., 2010), two evaluated the GBG (Dadakhodjaeva et al., 2019; Donaldson et al., 2017) and the remaining four evaluated various other game-based interventions (Hine et al., 2015; Hoff & Ervin, 2013; Mitchem et al., 2001; Wills, 2002). Although sample description did not serve as a moderator in the review, when considered individually, these studies have provided useful insights which highlight the value of looking at individual, as well as group data. Within the studies collated as part of the review, two patterns of responding were identified. Cases where the individual did not respond as well as the whole group, or as well as other individuals to an intervention, and cases where the individual responded just as well, if not better than the whole group to an intervention. Examples of these patterns of behaviour will be discussed next.

The first pattern of responding was evident in a study by Donaldson et al. (2017), who found that 3 out of 12 individual students monitored during the GBG did not respond favourably. Donaldson et al. proceeded to suggest that this method could be useful as a screening tool to identify students who do not respond to class-wide procedures. This could ensure behavioural resources are distributed appropriately based on data. In an evaluation of the CW-FIT intervention, Conklin et al., (2017) demonstrated incidences where individual behaviour did not improve to the same extent as the class group. The authors evaluated two outcomes across four whole-class groups (on-task behaviour and compliance) and four outcomes across multiple individuals within these four class groups (on-task behaviour, hand-raising, talking-out, out-of-seat). Improvements were evident in both individual and group behaviour across all study phases, however in the review in Chapter 2, slightly larger effect sizes were evident for positive group behaviour than positive individual behaviour, suggesting that some individuals may not have responded as well to the intervention as their whole corresponding class group. This reiterates the importance of examining individual behaviour in the context of the whole class group. Another example of this type of pattern of responding was in a study on the GBG by Dadakhodjaeva et al., (2019). Implementation of the GBG was found to produce improvements in academically engaged and disruptive behaviour across three classes and the three individuals monitored within these classes. However, the data were not as stable across individual students as for the class, and effect sizes were not as large for these individuals as for the whole class. This suggests that while individual students may respond to the intervention in improving their behaviour, they may not demonstrate improvements in line with the larger class group. These differential responses would have been masked had the researchers not collected data on the individuals. Notably, the three teachers involved rated the GBG neutrally rather than positively, with one teacher responding particularly negatively to the GBG. Although the authors do not speculate as to why these ratings were given, it is possible that teachers could not appreciate the strong positive effects of the intervention on the whole class due to the behaviour of the most disruptive students remaining unstable.

The second pattern of responding, where individual students responded more favourably to an intervention than their corresponding class group, was evident in a study by Hoff and Ervin (2013). This study evaluated the effects of a self-management game on whole class and individual student behaviour. Effects were slightly more potent across individual students than for the class group. This was evident in Tau effect sizes calculated as part of the review in Chapter 2 which produced Tau values of .42 and .59 for the effects on the whole class and individuals respectively. In this evaluation by Hoff and Ervin, class-wide behaviour was not hugely problematic at baseline, particularly in two of the three classrooms where disruptive behaviour occurred during less than 12% of intervals. Collecting data on the group and individual students allowed the authors to assert that the target students' behaviour was more 'in line' with that of their peers at the end of the study. This demonstrates another benefit of collecting data on individuals where class-wide behaviour is not an issue, as it allows for assessment of individual behaviour with the class-wide data as a reference point.

Tanol et al. (2010) also examined individual and group behaviour, this time in the context of the GBG-response cost and CBGG. The authors found that while both games were effective, the CBGG produced lower rates of rule violations across three of six individual target students (three targets in each of two classrooms). The other three targets students' behaviour was variable with high levels of overlap. The group data did not concern the whole class group but was an aggregate of the three target students' data in each class. These findings demonstrate further how monitoring of individual behaviour can produce rich experimental data which can enhance understanding of an intervention's effectiveness (or non-effectiveness as the case may be). There is the potential to identify students who require more intensive intervention, but equally the potential to identify instances where particularly disruptive students respond well to a universal intervention.

A study by Groves and Austin (2020) which was published since the review in Chapter 2 was conducted, examined the CBGG with three target individuals and three of their non-target peers in a Year 4 class in a Welsh school context. Specifically, the authors sought to examine the game with a known versus an unknown/mystery criterion for winning. The game was applied similarly to how the CBGG has been applied earlier in the current thesis (in Chapter 3), however rather than having the teacher award points contingent on good behaviour at the end of a specified interval, points were awarded at the end of every 2 min interval contingent on behaviour of the group being good for the *whole* preceding interval. Across target students and their non-target peers, both versions of the CBGG implemented were similarly effective, however the part-taking teachers preferred an unknown criterion. The CBGG produced reductions in problem behaviour across all six students who were monitored, however effects were larger for the target students, given they had a higher rate of problem behaviour at baseline. Collecting data on individual target and non-target students provided useful results, in that it allowed for demonstration that target students' behaviour could be more in line with their peers while taking part in the CBGG. It must be noted that although it provided useful insights, this method of data collection did not account for group effects, in what was a rather large class group (n = 33). This further supports the notion that examining both individual and group data may provide the richest evidence for the effectiveness of game-based classroom management interventions.

It should be noted that the review in Chapter 2 only identified two high-quality studies (i.e., studies meeting WWC design standards; WWC, 2017) on games like the CBGG while examining both individual and group behaviour (Conklin et al., 2017; Tanol et al., 2010). Conklin et al., conducted this in the context of the CW-FIT intervention package and although Tanol et al. (2010) report on collective data it is the collective behaviour of the three target students rather than whole class. Groves and Austin's (2020) study which was published after the review was conducted also meets the WWC standards, however compared target student data to that of their non-target peers. Research to date has therefore not examined whole class and individual student data in the context of the CBGG, without applying it as part of a larger package intervention (i.e., the CW-FIT). Therefore, one of the primary aims of the current study was to evaluate the effects of the CBGG on group and individual behaviour.

Game-based Interventions in Single-sex Environments

Chapter 2 identified the gap in the literature around the application of game-based interventions in single-sex classroom environments, with most studies identified reported a mixed-sex sample, or not reporting the gender of the sample. This may be partly due to the majority of the research on game-based classroom management interventions occurring in the USA, where co-education schools are more common than single-sex schools. For instance, in the 2014-15 school year, there were only 283 public single-sex schools in the USA (excluding special education, alternative education, vocational schools and juvenile justice facilities), and seven states had no single-sex schools (Mitchell et al., 2017). Given there were almost 100,000 public schools in the USA in 2014/15 (National Center for Education Statistics, 2019), this reflects a minimal percentage (<1%). In Ireland in 2019, 9.72% of primary schools were single-sex, and a further 2.29% were single-sex schools with mixed infant level classes (Central Statistics Office, 2020). Mixed infant-level classes refers to when the first two years of schooling (junior and senior infants) comprises mixed-gender classes, and students attend single-sex classes from 1st class upwards (i.e., the remaining 6 years of primary school). Boys attending single-sex schools made up 15.62% of the population of school-going boys in the country in 2019 (Central Statistics Office, 2020; this does not account for boys attending schools with mixed infants). The Growing Up in Ireland data, collected as part of the largest longitudinal study on child development in Ireland, has identified higher frequencies of misbehaviour among boys than girls (Williams et al., 2018) and in Irish primary schools, more punitive disciplinary measures (e.g., suspension) are used in single-sex boys' schools (Smyth & Quail, 2017). The relatively high percentage of single-sex schools combined with evidence of different levels of punishment in single-sex boys' schools, suggests a need to examine classroom management interventions in single-sex environments. It is also evident from Smyth and Quail's (2017) work, that boys' schools may need to place efforts into focusing on more positive behaviour management interventions, as has been recommended by the National Educational Welfare board (2008).

Although there has been a lack of research in single-sex environments, some research has looked at the impact of game-based classroom management interventions on males only as part of a larger class group. For example, Groves and Austin (2017) evaluated the CBGG in two formats with four male students attending a Pupil Referral Unit. There were only three other students in the class group, however the authors do not specify whether they were male or female and data were not collected on their behaviour. The target males were chosen based on data taken on their rates of disruptive behaviour. The CBGG was effective across the four target students in the reduction of varying types of disruptive behaviour. In another study by Groves and Austin (2020) which was discussed earlier in this Chapter, the three target students were male and the gender of the three non-target peers was not specified. In fact, the gender make-up of the school in which the study was conducted is not specified, so it is unclear whether the boys were chosen as targets from a class which was mixed- or single-sex. In the earlier discussed study by Tanol et al., (2010), the six chosen target students were male, and the CBGG was effective in targeting rule violations across all six. Again, the gender make-up of the whole class was not specified.

In summary, research on the use of a positive behavioural interventions among boys attending single-sex education in Ireland is warranted given the enduring popularity of single-sex education and the potentially high rates of punitive, reactive interventions in these settings. It is evident that the CBGG has yet to be investigated across a single-sex classroom, considering the impact of the game on individual and group behaviour. The dearth of research on the use of the CBGG in Irish primary schools more generally adds further to the need for more work in this area. It was therefore an aim of the current study to investigate the application of the CBGG in a single-sex, male-only classroom with a population deemed 'at-risk' of educational disadvantage. Furthermore, differentiating between target individual students' and the whole class response to the intervention is an important line of inquiry given the potential for non-responders, or indeed the potential for individuals to respond particularly well to the intervention. Another key aim of the current study was therefore to evaluate the CBGG with particularly disruptive students as identified by the class teacher.

Research Questions

The following research questions were proposed:

1. Is the CBGG effective in targeting academically engaged behaviour and disruptive behaviour across a whole class of boys in a single-sex boys', mainstream classroom environment?

2. Is the CBGG effective in targeting academically engaged behaviour and disruptive behaviour in two teacher-selected students displaying high rates of disruptive behaviour and/or low levels of engagement?
3. Does the teacher find the CBGG an acceptable and effective intervention for application with a male-only mainstream class?
4. Do students find the CBGG to be an acceptable intervention?

Method

Recruitment, Participants and Setting

Emails were sent to a number of primary schools briefly outlining the project and principals were asked to get in contact if they were interested in having one or more classes in their school take part. A principal of a single-sex boys' school contacted the primary researcher by email stating that a fourth-class teacher was interested, and a meeting was organised with this interested class teacher. Before this meeting took place, a representative from the school's board of management read about the study in a PLS and signed a consent form, indicating that the school was willing to allow researcher access for the course of the project (Appendix T1; T2). The project was outlined in detail to the teacher during the meeting and she was given relevant recruitment materials including a PLS and consent form (Appendix T3; T4). PLSs and consent/assent forms were then sent home in an information pack with each child in her class (Appendix T5-T7). Only students who returned fully completed parental consent forms and student assent forms participated in the study. All students could still take part in the game as the teacher decided to adopt it as her classroom management procedure during intervention phases. This simply meant that data were not collected by the researcher for those students who had not returned completed consent and assent forms.

The school setting was a single-sex boys' primary school in an urban area of Dublin. The school maintained DEIS status according to the Department of Education and Skills. The class teacher (Ms. Daly; pseudonym) was a 23-year-old female with two years of teaching experience. She had spent one of those two years in the part-taking school. The target class was at fourth class level and 18 students from the class took part. Fourth class is the sixth year of formal schooling in the Irish education system and is approximately equivalent to fourth grade in the US school system or Year Five in the English/Welsh school system. Students in fourth class tend to be approximately 9-10 years old. Ms. Daly chose two individual students for individual behaviour monitoring based on her evaluation of class behaviour. In choosing these students she was asked to consider those students who were most disruptive during class, while also considering their attendance, such that the students would be likely to be present during data collection sessions. The two male target students chosen for individual behaviour monitoring were Adam and Ben (pseudonyms).

Materials

The teacher was provided with two laminated copies of the class rules (Appendix U) and a scoreboard with space to write in individual team names (Appendix V). The teacher was given an Octopus watch which was pre-programmed to vibrate and remind her to scan the classroom intermittently. The watch was identical to that used in Study 2, Chapter 3. It had been planned to use the Fitbit Charge 2 as a prompting device in the current study, however this device would only allow a maximum of eight alarms/prompts. This was an issue as the prompting device needed to vibrate every 2 min for one hour in order to allow the teacher to begin the game at any stage she saw appropriate (i.e., the teacher may have been given the watch at 11.10am, but start to play the game at any stage between 11.10am and 11.40am). Therefore, the Octopus watch was used as a viable and readily available alternative. The teacher chose the prize based on what she had used with students before. It was also important that the prize was different to anything being used by the teacher in other classroom management procedures throughout the day. A ‘lucky dip’ was chosen as the reward given that a bag could be filled with a variety of small, cost-effective rewards (e.g., stationery, sweets). The teacher did consider whole class activities as potential rewards (e.g., a walk outside), however this would not be possible on days when all teams did not win the game. The teacher kept a checklist on her desk outlining the steps of the game, which was very similar to that used in Chapter 3 (Appendix W). Observers used the same checklist to carry out treatment integrity checks. Data was collected by observers with paper and pen (Appendix X) and interval changes were signalled via earphones.

Operational Definitions of Target Behaviours/Dependent Variables

Data were collected on academically engaged behaviour (AEB) and disruptive behaviour (DB). Both target behaviours are described in more detail here.

Academically Engaged Behaviour (AEB)

The operational definition used for AEB in this study was as follows:

"A student is academically engaged if they are oriented towards the academic task at the time of recording. This includes both active and passive engagement. Active engagement is occurring when a student is speaking to the teacher appropriately about the task, speaking to a peer appropriately about the task, writing during writing tasks or reading aloud when instructed to do so. Passive engagement is occurring when a student is oriented towards the academic task, but not actively engaged, for example, looking at the teacher, looking at the board, looking at their text-book. A student is not considered to be academically engaged if they are engaging in any of the listed disruptive behaviours at the time of recording. Exceptions to this is when a student is engaging in motor disruption which is compatible with the academic task e.g., the student is playing with a pencil, but simultaneously looking at the board."

Disruptive Behaviour (DB)

DB was monitored across three different categories: verbal disruption (VD), out-of-seat behaviour (OOS) and motor disruption (MD). These categories were chosen based on preliminary observations of the class and Ms. Daly's specific concerns. When analysing data each category of DB is considered separately as well as the composite variable of DB. Each category of disruption is defined here.

Verbal Disruption (VD). VD was defined as follows:

"A student is considered to be engaging in VD if they engage in a vocalisation not authorised by the teacher and unrelated to the work ongoing in the classroom, for example, whistling, singing, humming, shouting, cursing, talking to a peer about a topic unrelated to the work or talking to a peer when this has not been allowed".

Out-of-Seat Behaviour (OOS). OOS was defined as follows:

"OOS refers to any incident where a student's buttocks leaves their chair for >3 seconds without first obtaining permission, including standing at their place or standing and walking around the room. Swinging on the chair is also classified as out of seat behaviour i.e., one or more legs of the student's chair is being held off the floor. A student is not classified as being out of seat if they leave their seat and walk straight to the sink or bin and straight back. If the student converses with another student or

engages in another type of disruption while walking, then OOS is coded, and the type of disruption engaged in is also coded (e.g., if talking to a peer, VD is also coded)".

Motor Disruption (MD). MD was defined as follows:

"MD involves playing with objects (i.e., using one or two hands to manipulate an object) in a manner which is incompatible with the academic task (e.g., flicking paper across the room with a ruler), turning in one's chair away from the task (>3 seconds), leaving one's head on the desk, physically interacting with a peer in a manner which is incompatible with the task (e.g., hitting/touching a peer with a body part or object, playing a game with a peer etc.)".

Observation Procedures and Data Collection

Data were collected up to four times per week in the classroom during the period between 11.10am and 12pm. This was a period in between the students' short morning break and lunch time. Data collection sessions lasted 20 min. Class subject matter was not always the same during this period (i.e., some days it was mathematics, some days it was English etc.), however the teacher always presented a topic for approximately 10 min, followed by independent seatwork. The independent seatwork involved students being given an activity pertaining to the lesson which was ongoing which they would complete by themselves at their desk/seat. For example, if the lesson was English, and the teacher spent the first 10 min reading a book chapter, the independent seatwork may have involved students completing a worksheet about the chapter independently at their desks. Therefore, the lesson structure during this period was always similar. During baseline sessions, data collection commenced when independent seatwork commenced and during intervention sessions, data collection commenced when the teacher announced the beginning of the game. Data were collected via momentary time sampling and partial interval recording. An individual fixed method was used, as in previous studies in the current thesis. In this study, intervals were set to 10 s, with 5 s to record. A different student was observed during each 10 s interval, and a target student was observed every second interval, alternating between Adam and Ben. Therefore, in a 20 min observation period, each target was monitored for 20 intervals each and class-wide behaviour was monitored over 40 intervals. If one target student was absent, then the other target student who was present was observed for 40 intervals. On occasion, data collection sessions did not last the whole 20 min due to unforeseen circumstances such as a visitor coming to the class. However, this was rare, and a majority of the sessions lasted the full 20 min. There was one extended break in data collection in the middle of the first withdrawal phase. This break coincided with the classes' preparation for a school show for which the teacher had a significant workload and had to spend portions of the school day rehearsing in the school hall. This was the only non-scheduled break in data collection (i.e., that was not related to school holidays).

Data on AEB are presented as one variable whereas data on DB were collected across three categories. Each category of DB was considered separately, and the composite variable of DB was also considered. When creating the composite variable, an interval was considered to have DB present if *any* of the three types of DB occurred. Therefore, an interval was marked as containing DB whether one, two, or three types of DB occurred. The number of intervals marked with DB was then divided by the total number of intervals and multiplied by 100 to obtain a percentage of intervals with DB present.

Observer Training and Interobserver Agreement

Data collection was conducted by the student researcher (primary observer) and IOA data was collected by trained undergraduate and master's level psychology students (second observers). The majority of the IOA data was collected by the same observer, a third-year psychology student who was on a work-placement with the project supervisor. All observers took part in a training session where they learned about momentary time sampling, partial interval recording and frequency data collection methods. They were given a chance to ask questions and practice with videos before presenting to collect data. Second observers were blinded to study purpose and were not told what type of phase they were collecting data in (i.e., baseline phase, intervention phase) so that they would not have any expectations around what behaviour 'should' look like within that phase.

IOA data were collected on 25% of observation sessions overall and during at least 20% of the time in each study condition, as per WWC recommendations (WWC, 2017; 2020). It was collected at least once per phase for each outcome, however on one occasion, Ben was absent during an IOA observation session. Therefore, IOA was collected on 25% of occasions for the whole class, 25% of occasions for Adam and 21.43% of occasions for Ben. IOA was collected at least once per phase for the whole class and for Adam, and during three out of four phases for Ben (absent for the IOA observation session in the first intervention phase).

IOA for student behaviour was calculated using interval-by-interval agreement and dividing the number of agreements by the total number of observation sessions and multiplying by 100 to obtain a percentage. Table 4.1. outlines the overall mean IOA and range for each outcome and participant.

Table 4.1.

Mean IOA and Range Across Participants and Phases

Outcome	Whole Class	Adam	Ben
	M (range)	*M* (range)	*M* (range)
Verbal Disruption	96.92% (92.5-100%)	93.69% (85-100%)	96.67% (95-100%)
Out-of-Seat	98.79% (97.44-100%)	94.02% (89.47-100%)	98.25% (94.74-100%)
Motor Disruption	94.44% (90-97.44%)	91.46% (84.21-95%)	91.58% (90-97.44%)
Total Disruption	95.88% (90.83-98.29%)	93.02% (89.47-96.08%)	95.5% (95-96.49%)
Academically Engaged Behaviour	90.52% (79.49-97.73%)	88.02% (82.35-94.74%)	91.67% (80-100%)

Social Validity

Although it was intended that the teacher and students would complete social validity measures following the final day of data collection, data collection ceased unexpectedly in March 2020 following school closures due to Covid-19 restrictions. The teacher completed the Behaviour Intervention Rating Scale (BIRS; Elliott & Von Brock Treuting, 1991) remotely and forwarded it to the student researcher. The BIRS was identical to the version used in Chapter 3, Study 2. The teacher was asked to answer the questions with reference to the CBGG generally. An additional question asked the teacher whether she had additional comments to add about the game: “Do you have any further comments/feedback on the CBGG?”. Student social validity data were not collected due to government directed school closures amid the Covid-19 pandemic. Schools did not re-open until September 2020, therefore it was not possible to collect valid social validity data from students.

Experimental Design

A reversal/withdrawal design with phases A-B-A-B was used to evaluate the CBGG across the class and individual students taking part. The A phase refers to baseline and B refers to the CBGG with 2 min intervals between point opportunities. In other words, the teacher would scan the room to check for appropriate behaviour every 2 min and award points contingent on appropriate behaviour. It had been intended to probe the intervention further by including additional intervention phases with longer intervals between point opportunities (i.e., schedule thinning), however due to premature school closures, this was not possible. Chapter 5 will deal with schedule thinning during the CBGG in more detail.

Procedure

Baseline

During baseline, Ms. Daly proceeded with planned educational tasks and the CBGG was not in place. Ms. Daly used ClassDojo, a class-wide interactive system for awarding points to students, during this phase. She would award points to students for numerous reasons, such as for working quietly or answering a question correctly. This procedure was used throughout the whole school day and when students had built up enough individual points, they could swap them for a reward such as a homework pass. There was no interdependent group contingency procedure in place during this time and points awarded were at the teacher’s discretion, rather than at scheduled times throughout the day. Points could also be rescinded for disruptive behaviour. Ms. Daly agreed that she would not implement this strategy while the CBGG was in place, however she did continue to implement it throughout the day in intervention phases during periods where the CBGG was not in place. Given that it is commonplace for primary school teachers to implement a token-based strategy similar to Class Dojo, the implementation of this strategy during the baseline or intervention phase (when the CBGG was not in place), was not a major cause for concern. The CBGG would only be in place for 20 min during the school day and it was expected that the teacher would have alternative strategies for dealing with behaviour in place during other parts of the day.

During the last day of baseline data collection, the teacher was asked to privately implement the CBGG, without letting the children know that she was conducting behaviour checks at different intervals. This was employed for two reasons. The first was to allow Ms. Daly to get used to the vibrating prompts delivered by the Octopus watch every 2 min. It allowed her to practice awarding points when it prompted her, albeit without letting the students know that she was checking their behaviour. The second reason was to obtain a baseline estimate of how many points students could feasibly achieve. This number could then be used to set goals/a criterion during intervention phases. A similar approach to setting a criterion for the CBGG was employed in a study conducted by Ford (2017). Ford (2017) determined the points criterion by calculating the mean number of points acquired by classes during baseline, dividing it by two (to account for the team component) and adding 10-20% to make the criterion challenging but fair.

Teacher Training

Teacher training for implementing the CBBG was delivered across one after-school session during the first baseline phase. During this session, Ms. Daly was shown the behavioural data for the class and for the individual targets. The student researcher then took her through a brief training presentation (on PowerPoint) which outlined how to play the game, when the watch would vibrate and when and how to award points. Ms. Daly assisted in choosing a reinforcer. A preference assessment was due to be carried out however, Ms. Daly already used many of the potential reinforcers in her Class Dojo strategy. As well as this, many of the potential rewards she could think of either a) would require the whole class to win (e.g., a walk outside) or b) would require that only one team win (e.g., sit on cushions for the rest of the day). She therefore decided there were not enough alternatives which could be put on a preference assessment questionnaire and that a lucky dip would be the most appropriate and cost-effective reward for the group. The lucky dip was a bag filled with various edible treats and stationery such as pencils and rubbers.

Intervention: Caught Being Good Game

The CBGG was introduced to the class in the next class session following completion of baseline data collection. The teacher made it clear at the outset that during the CBGG, Class Dojo points would not be available, but CBGG points would be instead. The class was divided into four teams based on the existing group tables in the classroom. Each table served as a separate team. Each team could choose their own team name which was added to the scoreboard. Ms. Daly told the children about the game and told them that each time the game was played there would be 10 chances to earn a point if everyone on the team was following the class rules. These rules were as follows:

- I will remain in my seat with the four legs on the ground
- I will follow my teacher's instructions
- I will raise my hand before speaking
- I will try my best and focus on my tasks
- I will allow my classmates to try their best

The game was played daily for 20 min (one point every 2 min for 20 min = 10 points available in total). This was considered a reasonable amount of time which would fit well with the type of work to be completed at that time.

After explaining how the game was played and reviewing the class rules, Ms. Daly announced that the game had started. When she was prompted by the watch to check behaviour and award a point to teams following the class rules, she would do so on the scoreboard by placing a tally/point in the corresponding section. The goal for the 2 min version of the game was 5 points. This was calculated by taking the average points 'earned' by each team during the last day of baseline data collection (i.e., the session where Ms. Daly privately played the game) and adding 10%. The average amount of points earned across teams during the last day of baseline was 4.25. Adding 10% to this gave a total of 4.675 and this was rounded to a goal of 5 points, or 50% of the total points available during a game session. Prizes were awarded daily, immediately after the game session finished. Teams meeting or exceeding the 5-point goal were eligible to receive the prize. The team with the most points got the first pick from the lucky dip, followed by the team in second place and so on.

Treatment Integrity

The teacher checklist for game implementation for the current study was similar to that used in Chapter 3, Study 2, with the exception that the checklist incorporated daily rather than weekly prizes. There were 11 steps on the checklist for game completion. The teacher kept the checklist on her desk to serve as a prompt to complete each step. Treatment integrity data were collected during 100% of intervention data collection sessions by the student researcher.

Ms. Daly's average treatment integrity across the intervention phases was 95.45%. Her treatment integrity never dropped below 90.91% (range = 90.91-100%). All steps were implemented at least 75% of the time across intervention sessions. An overview of mean treatment integrity and range for each intervention phase is presented in Table 4.2.

Table 4.2.

Mean Rates and Range of Treatment Integrity during Intervention Phases for Ms. Daly

Phase	*M*	Range
CBGG 1	95.45%	90.91-100%
CBGG 2	95.45%	90.91-100%

Data Analysis

Data were analysed in the same way as in Chapter 3, taking guidance from the WWC design and evidence standards (WWC, 2017; 2020). Effect sizes were calculated using the Tau metric, as was conducted in previous studies. Again, Tarlow's (2017) recommendations were adhered to in the calculation of Tau via the Baseline-corrected Tau calculator (Tarlow, 2016). Weighted mean effect sizes were calculated for both outcome variables (AEB and DB) in the same way as in the review in Chapter 2 and the empirical work in Chapter 3.

Ethical Considerations

Previously discussed ethical principles and protocols were relevant to this study (see ethical considerations in Chapter 3), however two additional ethical amendments were developed pertaining to 1) the application of the CBGG in primary school settings and 2) the collection of data on individual as well as group behaviour. The original approved ethical protocol sought approval to carry out the research with secondary school students (i.e., adolescents) and to analyse data on group behaviour only. These amendments to the originally approved ethical protocol were submitted to DCU REC and subsequently approved before recruitment for this study commenced. The key points and risks addressed in these amendments will be discussed here.

The CBGG with Primary-school Aged Children

The first amendment was concerned with extending recruitment to primary schools. By engaging with a younger population in research activities, some greater risks are present given their vulnerability. The risks identified were considered not to outweigh those present in every-day school situations, however. For example, a group contingency intervention applied during the research project involves teams, goals, and rewards. This may be very similar to other classroom management strategies used by teachers such as token boards and star charts. Students were also likely to have worked in team situations in the school or sports/games settings.

Protocols around children becoming distressed remained similar to the previous studies conducted in secondary schools. Should any child have felt uncomfortable during the research project, due to engagement in the intervention or due to being observed by researchers, the student may have raised this with the teacher. The teacher would be informed they may stop the intervention in this case and deal with the situation in line with school policy. This may have involved withdrawing the intervention altogether or referring the child to other supports within the school setting. Observers would cease to observe the child during observation sessions. The research team made themselves familiar with school policies before observations began and second observers who conducted IOA observations were made aware of relevant ethical considerations during IOA training. No incidents of student or teacher distress arose during this study.

Individual Data Collection

A second ethical amendment was submitted to DCU REC outlining the protocols and ethical considerations around collecting data on individual behaviour as well as group behaviour during the research project. The proposed change outlined that data may be analysed at the individual level as well as the group level. The change was not concerned with how and what data were to be collected, but rather how the data were used and potentially how these data informed the intervention. The individual students whose data would be considered at the individual level would be selected in conjunction with the class teacher, based on the student’s engagement in the target behaviours for which data were to be collected.

This is iterative research and the studies reported in the previous chapter, as well as developments in the field, meant that a new research question was developed which necessitated this amendment. Research has shown that although a group contingency intervention may appear to be effective at the whole group level, individual students who display high levels of disruption or very low levels of engagement may not respond as positive to intervention (e.g., Dadakhodjaeva et al., 2019). When collecting and analysing group data, it is impossible to make informed decisions as to whether the most disruptive students in the class are benefitting from the intervention even if the class responds well. Through working closely with teachers in previous iterations (Chapter 3) of the game, it was evident that they felt that perhaps the most disruptive students are not always responding well. By identifying students who may not be responding to interventions at the group level, these students may be referred for more specialised intervention services within the school. These referrals would not take place within the scope of this project, but the collection of individual data may add to the literature on the monitoring of individual behaviour during group contingency interventions as a potential screening tool (Donaldson et al., 2017).

In keeping with previous considerations around confidentiality and anonymity, data collected on individual students remained anonymous and confidential. During data collection, students' names were not written on the sheets and instead they were assigned a number or letter (e.g., Target 1, Target 2). When reporting the data, names of individual students were changed to pseudonyms.

Results

WWC Design Standards

The CBGG (i.e., the independent variable) was systematically manipulated throughout the study, meaning the first of the WWC design standards was met (WWC, 2017; 2020). IOA data were collected at least 20% of the time overall and at least once per phase for whole class data and for Adam's data. Ben was absent for one data collection session where a second observer was present, so IOA data were not collected for Ben during that phase. Therefore, the study overall meets this criterion, however Ben's data do not. All IOA data met minimum thresholds of 80% on average. There were at least three attempts to demonstrate an intervention effect for each participant and there were at least three data points in each phase, but not at least five. Therefore, the study overall meets the WWC design standards with reservations.

Visual Analyses

Class-wide Effects

Class-wide AEB and DB data are presented across Figures 4.1. and 4.2.

Academically Engaged Behaviour. AEB across the whole class group was variable during baseline, and although it was high initially, it decreased towards the middle of the phase, and increased again towards the end of the phase (*M* = 76.33%, range = 66.67-87.5%). When the CBGG was introduced, AEB increased immediately and moderately to a mean of 92.5% of intervals (range = 90-97.5%). There was no overlap in AEB between this intervention phase and the preceding baseline phase, despite high levels of AEB at some points during baseline.

When the game was withdrawn, AEB immediately decreased and remained stable across the entire withdrawal phase (*M* = 76.67%, range = 70-82.5%). Again, there was no overlap between this withdrawal phase and the preceding intervention phase and behaviour was similar to that in the middle of the initial baseline phase. When the CBGG was reinstated, AEB increased immediately and substantially (*M* = 93.3%, range = 90-97.5%). There was no overlap between this phase and the preceding withdrawal phase and behaviour was very similar to the first CBGG phase. AEB was very stable across this phase.

Disruptive Behaviour.

Verbal Disruption. At baseline, the whole class engaged in a moderate level of VD (*M*= 15.83%, range = 7.5-22.5%). VD was variable across the phase, with very low to moderate levels reflected in the data. When the CBGG was introduced, VD initially remained at a level similar to the last point in the baseline phase, and then decreased across two data points, before increasing very slightly again towards the end of the phase (*M* = 5.63%, range = 2.5-7.5%). Data were much more stable across this phase compared to baseline with little variability observed, however two data points overlapped with the baseline phase.

When the CBGG was withdrawn, VD did not change substantially. It increased slightly at the beginning of the phase but occurred at a low level generally (*M* = 6.67%, range = 2.5-10%). In the final CBGG phase, VD did not occur at all (*M* = 0%, range = 0%). This reflects a substantial decrease when compared to both baseline phases and even the first intervention phase.

Out-of-seat Behaviour. OOS was not the main behavioural issue in this class and did not occur at a very high rate at baseline (*M* = 3.83%, range = 0-10%). Nonetheless, on introduction of the CBGG, there was a noticeable and stable decrease in OOS (*M* = 1.25%, range = 0-2.5%). When the game was withdrawn over three sessions, there was essentially no change in this behaviour and OOS remained low and stable across the phase (*M* = 1.67%, range = 0-5%). In the final CBGG phase, OOS did not occur at all (*M* = 0%, range = 0%).

Motor Disruption. MD was variable during baseline (*M* = 10.83%, range = 2.5-16.67%). When the CBGG was introduced, MD decreased and occurred at a low level across the phase (*M* = 4.38%, range = 0-7.5%). There was some overlap with the initial baseline phase. When the game was withdrawn, there was a small increase in overall level of MD, however it remained at a low rate across the phase (*M* = 8.33%, range = 5-12.5%). When the game was reinstated, this low rate of MD continued across the four data points (*M* = 4.89%, range = 0-7.5%).

Disruptive Behaviour Composite Variable. DB across the whole class was variable at baseline (*M* = 27.17%, range = 12.5- 35%). Despite high levels across the initial part of the phase, DB decreased during the final data point. DB remained at this low level when the CBGG was introduced and was stable across this phase (*M* = 10.63%, range = 7.5-17.5%). One data point overlapped with the initial baseline phase.

When the CBGG was withdrawn, DB did not increase immediately, but did increase in level overall (*M* = 15%, range = 12.5-17.5%). This increase in level was not to the same level as during the initial baseline phase however, which had been high over four of five data points (*M* = 27.17%). Finally, when the CBGG was reinstated, DB decreased immediately and moderately and remained at a low level across the phase (*M* = 4.89%, range = 0-7.5%). There was no overlap between this phase and the preceding withdrawal phase.

Figure 4.1.

Percentage of Intervals with AEB and DB Across Study Phases for Ms. Daly's Class Group

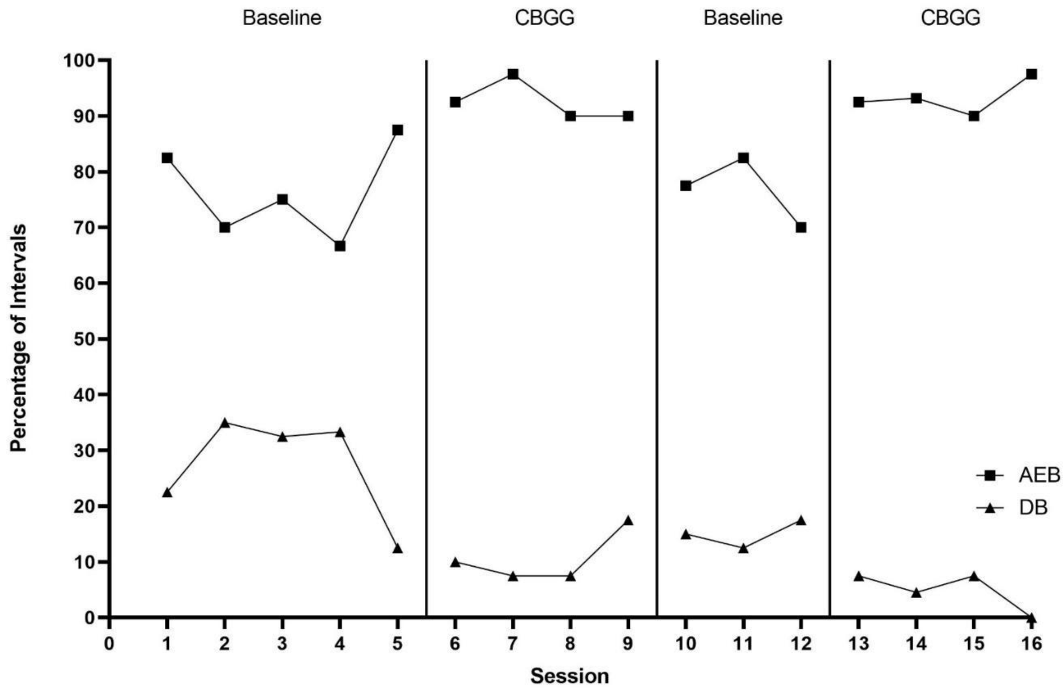
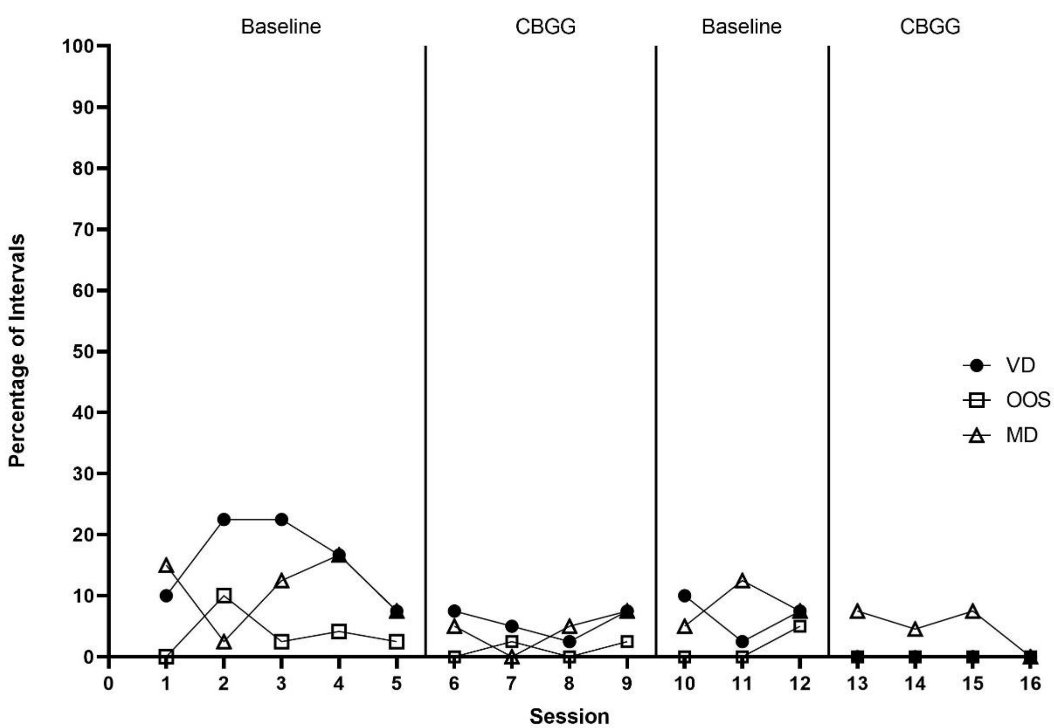


Figure 4.2.

Percentage of Intervals with VD, OOS and MD Across Study Phases for Ms. Daly's Class Group



Target Students

Adam and Ben's AEB and DB data are presented across Figures 4.3. to 4.6.

Academically Engaged Behaviour. Adam's AEB was slightly variable across the initial baseline phase (*M* = 76.17%, range = 70-83.33%), while occurring at a level similar to that of the whole class. When the CBGG was introduced, there was an immediate increase in his AEB, however it decreased again during the second session in this phase. There was an increase towards the end of the phase, with his AEB as high as 100% during one data point (*M* = 88.75%, range = 75-100%). His data were less stable than whole class data, but overall, the levels remained quite similar. One data point in this phase overlapped with the baseline phase.

When the CBGG was withdrawn, Adam's AEB decreased immediately and continued to decrease across the phase, until it was extremely low at the end of the phase (*M* = 55%, range =25-80%). When compared to whole class data, Adam's AEB during this phase was substantially worse. When the CBGG was introduced for the second time, Adam's AEB increased to 100% immediately. However, his AEB decreased to a very low level on the second day of data collection during this phase. It increased again towards the end of the phase (*M* = 71.69%, range = 11.76-100%). Overall, Adam's AEB was much more variable than that of the class. Adam displayed two extreme lows in AEB; once during the withdrawal phase (data point 12) and once during the final CBGG phase (data point 14). These extremes were not evident in whole class data (presented in Figure 4.3.).

Figure 4.4. outlines Ben's AEB across phases. Ben's level of AEB was low at baseline, occurring at a lower level than the whole class. He was absent for one data collection session during baseline, and his AEB did not exceed 65% (*M* = 50%, range = 30-65%). When the CBGG was put in place, Ben's AEB increased immediately and remained high across the three data points for which he was present (*M* = 86.67%, range = 80-90%). His behaviour during this phase was much more in line with the whole class, and during two data points (8 and 9; Figure 4.4), his behaviour was identical to that of the whole class. There was no overlap between this phase and the preceding baseline phase.

There was an immediate decrease in Ben's AEB when the intervention was withdrawn across three sessions (*M* = 61.67%, range = 55-75%). Initially, the decrease was in line with the whole class, however Ben's AEB deteriorated further and during the final two sessions occurred at a much lower rate than the whole class. An immediate and large increase was observed when the CBGG was reinstated for the final four sessions (*M* = 88.69%, range = 80-95%). Again, there was no overlap here with the preceding withdrawal phase and Ben's behaviour was generally similar to that of the whole class.

Disruptive Behaviour.

Verbal Disruption. Adam's level of VD was very high initially before decreasing and stabilising across the baseline phase (Figure 4.5.). VD remained quite high (*M* = 36.67%, range = 25-60%) across the baseline phase and occurred at a higher rate than the whole class generally. When the CBGG was introduced, his VD decreased immediately and substantially (*M* = 11.25%, range = 0-20%), despite still occurring at a rate higher than the whole class. There was no overlap here with the initial baseline phase despite some variability being present in the data.

Upon the game's withdrawal, Adam's VD was low initially, but increased substantially across the final two data points to 50% (*M* = 38.33%, range = 15-50%), bringing behaviour in line with initial baseline levels. An immediate decrease was apparent when the CBGG was reimplemented (*M* = 7.13%, range = 0-23.53%), although there was one data point in this phase with a high level of VD (data point 14; Figure 4.5.). This high level of VD corresponds with a very low level of AEB during the same data point.

VD was not the most concerning of the three types of DB for Ben, however it still occurred periodically at baseline (*M* = 12.08%, range = 8.33-15%), albeit at a lower level than the whole class. The behaviour was low and stable. When the CBGG was introduced, Ben's VD increased and was variable across the three sessions for which he was present, ranging from very low VD (0%) to moderately high VD (*M* = 13.33%, range = 0-20%).

When the game was withdrawn, Ben's VD increased further and continued at an increasing rate across the phase (*M* = 36.67%, range = 20-55%). When the game was reinstated, Ben's VD decreased to almost 0% across the four data points in the intervention phase (*M* = 2.5%, range = 0-5%).

Out-of-seat Behaviour. OOS occurred at a low and stable rate for Adam across the baseline phase (*M* = 6.67%, range = 0-10%) and did not occur at a much higher rate than OOS across the whole class. Despite the initial low levels, OOS still decreased to 0% for the first three sessions when the CBGG was introduced and occurred during 5% of intervals during the final data point in this phase (M = 1.25%, range = 0-5%).

When the game was first withdrawn, Adam's OOS increased immediately to 10% before increasing further towards the end of the phase (M = 15%, range = 10-25%), occurring at a higher rate than that of the class. When the CBGG was put back in place in the final phase, OOS did not occur at all (0%). This reflected a large decrease when compared to the preceding withdrawal phase and brought Adam's OOS in line with whole class OOS.

OOS was not an issue for Ben at baseline and occurred at a very low and stable rate (M = 2.5%, range = 0-10%), similar to that of the whole class. When the CBGG was introduced, OOS similarly remained low and stable (M = 1.67%, range = 0-5%). There was a slight increase when the game was first withdrawn, however there was a decreasing trend across the phase and OOS decreased to 0% again by the end of the phase (M = 5%, range = 0-10%). When the game was put back in place in the final phase, OOS remained low and stable for Ben across the phase (M = 2.5%, range= 0-5%).

Other Motor Disruption. Adam's MD was very similar to that of the whole class during baseline and occurred at a stable rate across the phase (M = 10.33%, range = 5-16.67%). When the CBGG was introduced, there was no substantial change in Adam's MD (M = 8.75%, range = 0-15%) compared to baseline levels. There was also no change relative to the notable change observed in the whole class MD data.

When the game was withdrawn, Adam's MD was initially similar to the first two phases, however increased across the phase to a higher level than the whole class (M = 23.33%, range = 15-30%). MD decreased to 0% when the CBGG was reinstated, however increased to a very high rate during data point 14 (52.94%; Figure 4.5.), corresponding with a very low rate of AEB during this session. Following this data point, MD returned to a lower level at the end of the phase (M = 15.74%, range = 0-52.94%).

Ben engaged in a high rate of MD during baseline (M = 48.33%, range = 30-58.33%) which was substantially higher than whole class MD (M = 10.83, range = 2.5-16.67; see Figure 4.4.). When the CBGG was put in place, there was a large decrease which remained stable across the three data points for which Ben was present (M = 6.67%, range = 5-10%). This brought Ben's behaviour much more in line with the class and there was no overlap here with the initial baseline phase. There was an immediate increase when the CBGG was withdrawn, however data were slightly variable during this phase (*M* = 20%, range = 10-30%). When the game was reinstated, MD was initially high, but decreased during the second data point and remained low across this phase (*M* = 11.32%, range = 5.26-20%).

Disruptive Behaviour Composite Variable. Adam engaged in a high level of DB at baseline, although it was variable across the phase (*M* = 46%, range = 35-70%). Generally, his level of DB remained higher than that of the whole class (see Figure 4.3.). When the CBGG was put in place, DB decreased immediately and substantially, although it remained variable across the phase (*M* = 17.5%, range = 0-30%). There was no overlap here with the initial baseline phase.

When the CBGG was withdrawn, Adam's DB increased immediately and continued to increase at a steep rate across the three data points in this phase (*M* = 58.33%, range = 35-80%). When compared with whole class data, Adam's DB occurred at a much higher rate with a steep increasing trend not observed across the class. This phase did not overlap with the preceding intervention phase. In the final CBGG phase, DB was very low initially, but increased to a very high level during the second data point (data point 14; Figure 4.3.). It then decreased during the final two data points in the phase (*M* = 21.4%, range = 0- 70.59%). Despite behaviour being very variable in this phase, DB occurred at a lower level generally than either of the baseline phases.

Ben's rate of DB was high and somewhat variable at baseline (*M* = 54.58%, range = 35-65%) and occurred at a substantially higher level than the whole class DB. DB decreased immediately and substantially when the CBGG was put in place, however remained variable across the phase (*M* = 20%, range = 5-30%). Despite this variability, the data did not overlap with the initial baseline phase.

When the CBGG was withdrawn, there was a large increase in DB, with an increasing trend across the phase (*M* = 48.33%, range = 40-60%). During this phase, Ben's DB was again substantially higher than whole class DB. There was a large and immediate decrease when the CBGG was reinstated (*M* = 15.07%, range = 5.26-25%). There was no overlap between this intervention phase and the preceding withdrawal phase.

Figure 4.3.

Percentage of Intervals with AEB and DB Across Study Phases for Adam Superimposed over Whole Class Data for Comparative Purposes

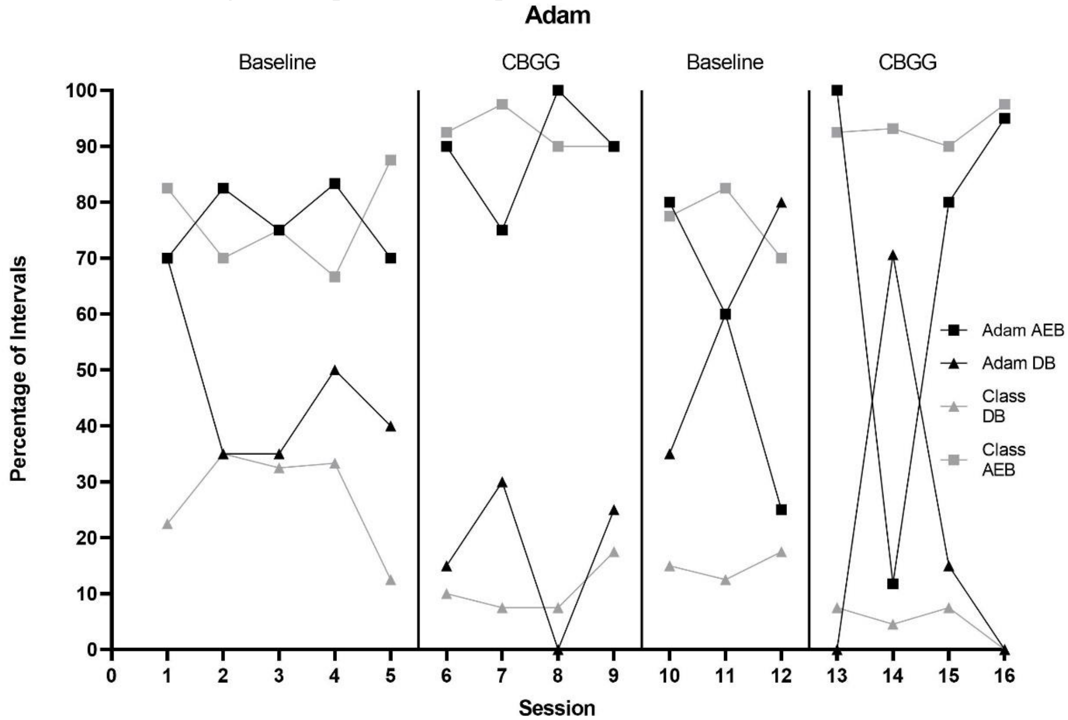


Figure 4.4.

Percentage of Intervals with AEB and DB Across Study Phases for Ben Superimposed over Whole Class Data for Comparative Purposes

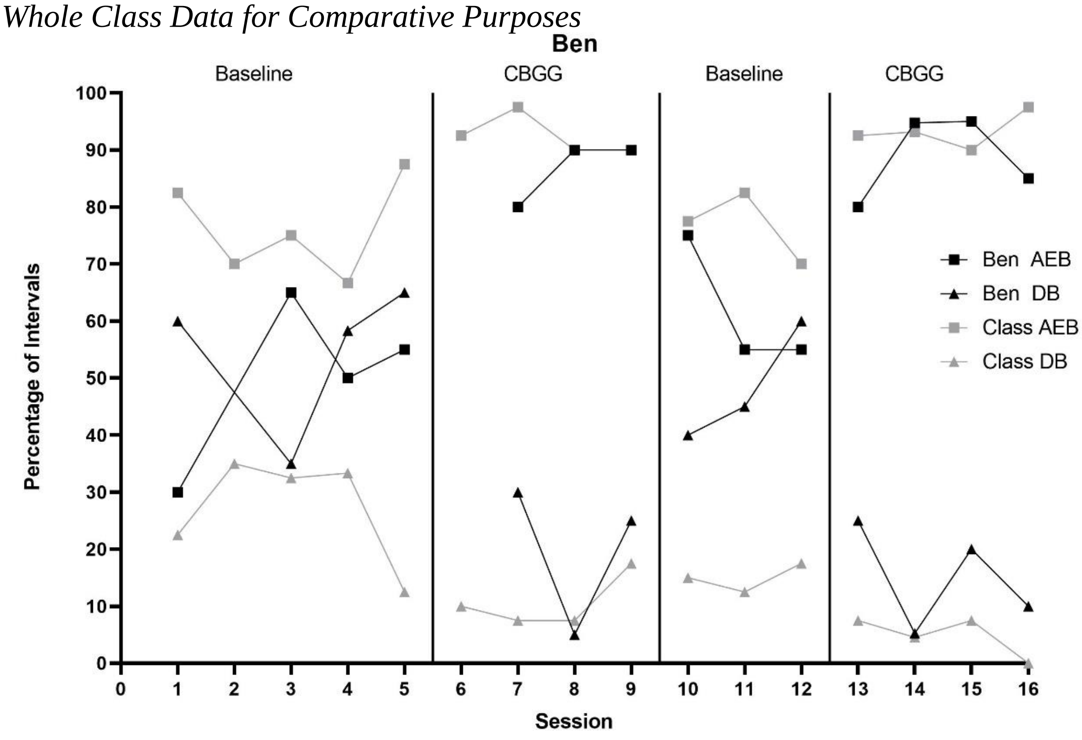


Figure 4.5.

Percentage of Intervals with VD, OOS and MD Across Study Phases for Adam

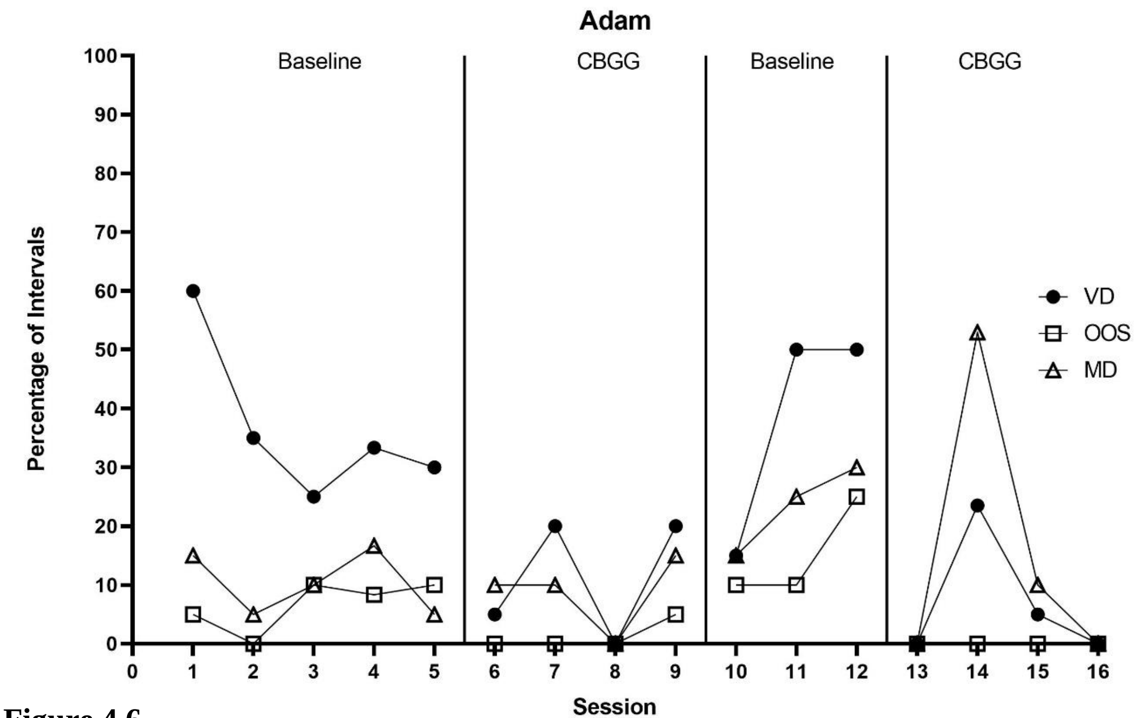
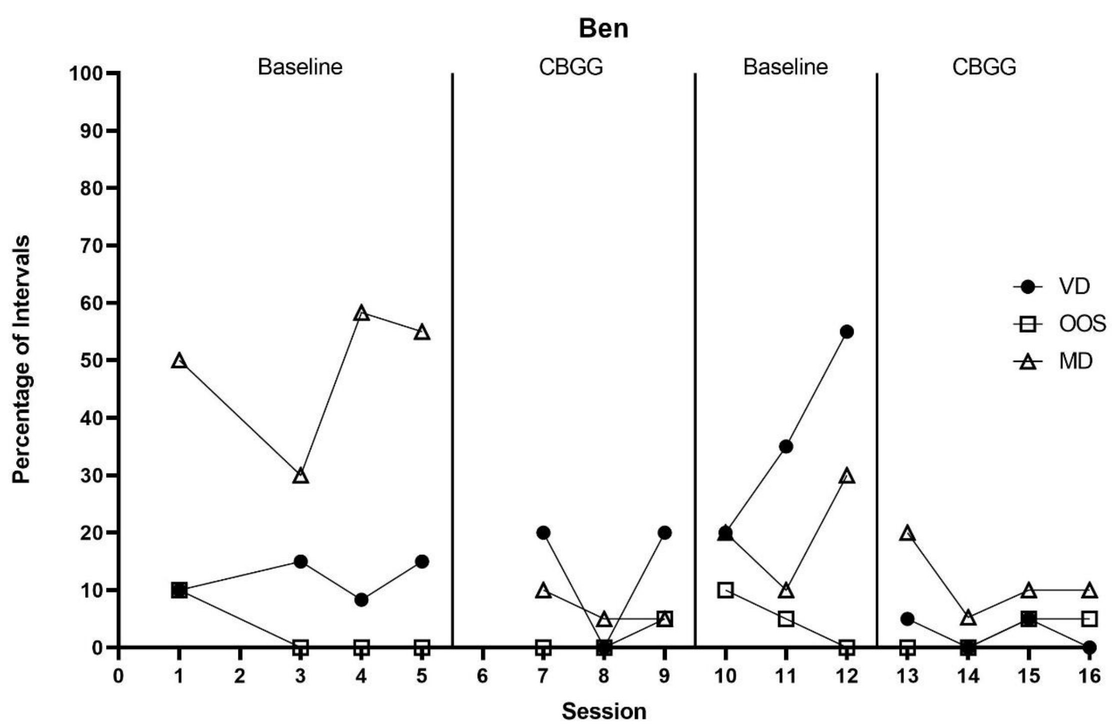


Figure 4.6.

Percentage of Intervals with VD, OOS and MD Across Study Phases for Ben



Effect Sizes for Behaviour Data

Tau effect sizes for the whole class, Adam and Ben's behaviour are presented in Table 4.3. Effect sizes are presented for each phase change (i.e., the change from phase A1-B1 and from phase A2-B2), as well as for the weighted mean across the two phase changes.

Table 4.3.

T*au Effect Sizes Across Phase Changes for the Whole Class and Target Students*

Outcome	Tau A1-B1	Tau A2-B2	Weighted Mean Tau
Whole Class			
Academically Engaged Behaviour	.756	.756	.76
Verbal Disruption	-.712	-.894	-.83
Out-of-seat Behaviour	-.43	-.471	-.45
Other Motor Disruption	-.499	-.408	-.46
Total Disruptive Behaviour	-.68	-.775	-.73
Adam			
Academically Engaged Behaviour	.584	.323	.49
Verbal Disruption	-.756	-.662	-.72
Out-of-seat Behaviour	-.592	-.926	-.85
Other Motor Disruption	-.12	-.387	-.25
Total Disruptive Behaviour	-.756	-.516	-.68
Ben			
Academically Engaged Behaviour	.775	.775	.78
Verbal Disruption	.265	-.795	-.49
Out-of-seat Behaviour	0	-.298	-.16
Other Motor Disruption	-.775	-.49	-.68
Total Disruptive Behaviour	-.756	-.756	-.76

Whole Class

For the whole class, Tau values for individual phase changes for AEB were both .756 which constitutes a large effect. The weighted mean for AEB was also large (.76). Tau values for subtypes of DB ranged from -.408 to -.894. The weighted mean Tau values ranged from moderate to very large for types of DB across the whole class (-.45 to -.83). The Tau values for the composite DB variables across phase changes were both large, as was the weighted mean (-.73).

Adam

For Adam, Tau values for AEB were .584 and .323. The weighted mean was moderate (.49). Tau values for individual phase changes for the three types of DB ranged from -.12 to -.926. The weighted mean Tau values for DB ranged from moderate to very large (-.25 to -.85). Tau values for the composite DB variable were moderate and large for each phase change respectively, and the weighted mean was large (Tau = -.68).

Ben

The Tau values for Ben's AEB for both phase changes was .775. The weighted mean Tau for Ben's AEB reflected a large change (.78). Tau values for individual phase changes for DB ranged from .265 to -.795. Ben's VD increased slightly from phase A1-B1, therefore the Tau value was positive. The weighted mean Tau values ranged from small to large for sub-types of DB (-.16 to -.68). For the composite DB variable, Tau values were large (Tau = -.756) for each phase change and the weighted mean was also large (Tau = -.76).

Social Validity

Teacher Social Validity

Ms Daly provided largely positive ratings for the CBGG on the BIRS (Elliott & Von Brock Treuting, 1991). She scored the game 82 out of a possible 90 on the acceptability subscale (*M* = 5.47). On the effectiveness subscale, she scored the game 23 out of a possible 42. Her mean rating on this scale was 3.29 (range = 3-4). Ms. Daly's rating on the efficiency subscale was 9 out of a possible total of 12. Her mean rating on this scale was 4.5 (range = 4-5). The items with which the teacher slightly disagreed were as follows: "The intervention will produce a lasting improvement in the students' behaviour"; "The students' behaviour will remain at an improved level even after the intervention is discontinued"; "Using the intervention should not only improve the students' behaviour in the classroom, but also in other settings (e.g., other classrooms, home)"; "The intervention produced enough improvement in the students' behaviour so the behaviour no longer is a problem in the classroom"; "Other behaviours related to the problem behaviour(s) were also improved by the intervention".

In the general feedback section where the teacher could write her own comments, she responded very positively to the game. The full response was as follows:

> *"I think this was a great initiative to introduce to the classroom, as it made them more aware of the classroom rules and I could tell that they were more conscious of following the rules while playing the game. I love how the good behavior is rewarded and the kids who are always good were being praised for their behavior, as sometimes this can go unrecognized in the busy classroom setting. It was also great how the class were working as teams as I felt like this encouraged the good behavior more, as nobody wanted to let their team down. I did find that when the game ended the class tended to think they did not have to follow the rules anymore as the game had ended, so I think that some sort or initiative could be set out to try and discourage this behaviour. However I thought this was a great initiative which the kids in my class loved, and I will definitely use this in the future".*

Student Social Validity

As noted previously, there was no opportunity to collect student social validity measures during this study. However, anecdotally, the student researcher can report that students would often ask the teacher 'Are we playing the game today?' and were disappointed on days when the game was not being played. Ms. Daly also alluded to the fact that the students enjoyed the game in her social validity considerations.

Discussion

The current study aimed to examine the effectiveness of the CBGG in a single-sex boys' mainstream classroom setting while investigating the effects of the intervention on both the whole class group and on two individual students. One fourth-class group (n = 18) and their teacher took part in the study, which spanned across four phases: Baseline, CBGG, Withdrawal, CBGG. Across the whole class, the CBGG was particularly effective in targeting AEB, and demonstrated effectiveness in targeting DB generally, especially VD. The CBGG was also effective at targeting AEB and DB in two target students, although to varying degrees. The teacher rated the game positively and stated that she would use it in future.

CBGG Effectiveness Across the Whole Class

Upon visual analysis of the whole class data, it was evident that the CBGG was highly effective in targeting AEB across this group of students. There was an increase in AEB each time the CBGG was implemented and there were no data points in the baseline phases which overlapped with the CBGG phases. Large effect sizes were identified for each phase contrast and the weighted mean effect size was also large (.76). This finding is in line with much other research demonstrating the efficacy of the CBGG with similar age groups. Wright and McCurdy (2012) demonstrated the efficacy of the CBGG in targeting on-task behaviour with a fourth-grade class which were a similar age to the fourth-class in this study. Wright and McCurdy recruited a mixed-sex class (10 male, 10 female), whereas the current study focused on a single-sex setting.

Levels of DB across the class during CBGG phases were consistently lower than during the baseline phases in the current study. The CBGG produced very large effect sizes in targeting VD and large effect sizes for DB overall. In the final CBGG phase, VD did not occur at all during observation sessions. Wright and McCurdy (2012) found that the CBGG produced a decrease in DB in the fourth-grade group recruited for their study, and these findings are in line with the current findings. Wahl et al., (2016) looked at the effects of the CBGG and GBG on off-task verbal behaviour with younger schoolchildren, and also found that in a class where verbal disruption was an issue, the CBGG and GBG produced meaningful decreases in behaviour. In examining three types of DB as well as overall DB, the current study contributes more substantial data to the literature. It allows for objective analysis of which type of DB is the most problematic and therefore needs most focus. Although Wahl et al., (2016) collected data on different types of DB, these were not graphed for visual analysis, nor were data collected on individual students.

The current findings also provide further support for the use of the CBGG as a potential alternative to the GBG across a whole class setting. In considering the CBGG as a viable alternative, it is useful to directly compare these findings to recent studies examining the traditional GBG. Dadakhodjaeva et al., (2019) conducted one of the most recent evaluations of the traditional GBG and found that it produced large to very large effect sizes on AEB and DB across three kindergarten classrooms. The Tau-U calculator (Vannest et al., 2016) was used by Dadakhodjaeva et al., (2019) to calculate these effects. Large effect sizes were detected in the current study for the effect of the CBGG on AEB and DB, however these effect sizes were calculated using a more conservative method (Tarlow, 2016, 2017) than that used by Dadakhodjaeva and colleagues, meaning a direct comparison cannot be made. Had the less conservative Tau-U calculator been used in the current study to calculate Tau scores, more 'very large' effect sizes may have been detected, bringing results in line with Dadakhodjaeva and colleagues. This is a promising finding given the focus on positive behaviour supports in schools and the move away from punitive, reactive strategies, which are commonplace in single-sex boys' primary schools in Ireland (Smyth & Quail, 2017). Overall, the findings here support the effectiveness of the CBGG with a single-sex mainstream class, with large to very large effects present for VD, overall DB and AEB and moderate effects present for MD and OOS.

CBGG Effectiveness Across Individual Target Students

Academically engaged behaviour was variable for both target students at baseline. Adam's rate of AEB was reasonably high on average (>70%). Adam's AEB remained highly variable across the course of the study, however, in general, the CBGG produced moderate improvements. This is reflected in an effect size which suggests the CBGG was moderately effective in targeting AEB for Adam (Weighted mean Tau = .49). One extreme low was detected in Adam's AEB (data point 14; Figure 4.3.), which the primary observer could not attribute to anything in particular. It was noted that at the beginning of this session, Adam stated that he 'didn't care' about the game and he was subsequently placed on an independent contingency by the teacher when he caused his team to miss opportunities to gain points. He needed to earn two points independently to get back on his team, however he did not achieve this. As it happened, his team did not earn enough points for a prize during this session, even without his influence. It was decided that his behaviour would be monitored closely over the following days and that if low AEB and high DB continued, a preference assessment would be carried out with Adam to ensure the prize on offer was reinforcing for him. Interestingly, Adam's AEB increased during the following observation session (two days later) and AEB remained high into the final data point. When comparing Adam's AEB to that of the whole class, it was evident that at times it was similar, however more often it was substantially lower. This was particularly evident during the withdrawal phase, where Adam's AEB substantially decreased over three data points. The key differences between Adam's AEB and that of the whole class related to variability and extremities in behaviour. Adam's AEB was highly variable, even during CBGG phases and extreme lows occurred at different points.

Ben's AEB was substantially lower than Adam's and the whole class at baseline (M = 50%). When the game was put in place however, Ben's AEB was much more stable. Immediate and substantial increases occurred when CBGG phases were introduced and there was an immediate decrease when the game was withdrawn. There were no data points in CBGG phases which overlapped with the baseline or withdrawal phases. This is reflected in a large weighted mean effect size (Tau = .78). Ben's AEB was more similar to that of the whole class during intervention phases. This is particularly evident during data points 8, 9, 14 and 15 (Figure 4.4.). This suggests that the CBGG is a useful intervention in bringing an individual student's behaviour more in line with their well-behaved peers.

The target students displayed tendencies towards differing types of DB at baseline. Adam engaged in a high rate of VD whereas MD was Ben's most prominent behaviour issue. The CBGG was effective in targeting both students' most prominent behaviour issues, while also being effective in targeting other issues and in targeting DB as a composite variable. Adam's overall DB during CBGG phases only overlapped with baseline phases during one data point (data point 14, which was discussed earlier), and there was no overlap at all observed for Ben. There was some overlap observed when examining individual types of DB, however this was to be expected, particularly for the sub-types which occurred at a low rate. Adam's VD, which was problematic at baseline, decreased substantially when the CBGG was in place versus baseline phases, with a large weighted mean effect size overall (-.72). A very large weighted mean effect size was also observed for the effect of the CBGG on OOS (-.85). Ben's MD, which occurred during a mean of 48.33% of intervals at baseline, decreased substantially when the CBGG was implemented, particularly during the first iteration of the game. Overall, the CBGG produced a large weighted mean effect size for Ben's MD (Tau = -.68). Both target students displayed levels of DB much higher than that of the group in general at baseline and during the withdrawal phase. Although their behaviour remained more variable than class behaviour during CBGG implementation, it approximated group behaviour more closely.

Dadakhodjaeva et al., (2019) investigated the effects of the GBG on whole class and individual student behaviour, and found that while the GBG produced large to very large effects on AEB and DB for the classes as wholes, effect sizes were moderate to large for individual target students. Effect sizes for all target students' AEB was reported as moderate, whereas in the current study, effect size for AEB was large for the whole class and for Ben, and moderate only for Adam. For the purposes of the review in Chapter 2, effect sizes for the Dadkhodjaeva study were calculated using Tarlow's recommendations. Use of this effect size metric produced a negative effect size for the impact of the GBG on one of the target student's AEB as baseline correction was recommended and carried out. This does not necessarily suggest that the GBG produced a negative impact on AEB, however it does suggest that the particular target student was demonstrating an improvement at baseline. Target students in the current study were chosen in conjunction with the teacher, however Dadakhodjaeva et al. do not report how targets in their study were chosen. Overall, it is evident that the CBGG examined in the current study compares favourably to recent implementations of the GBG (i.e., Dadakhodjaeva et al., 2019). Clearer conclusions may be drawn from comparing the CBGG with the GBG in the same class group. This was not the focus of the current study and was outside of its scope.

Groves and Austin (2020) demonstrated the efficacy of the CBGG with a cohort of students of a similar age profile to the current sample. In their study, the CBGG was effective in reducing problem behaviour across three teacher-selected targets students, both with and without a known criterion for winning. The current findings therefore align closely with these results, in that they provide further evidence for the CBGG with particularly disruptive students attending mid to late primary school grades. One key difference between the current study and that of Groves and Austin, is that rather awarding points contingent on good behaviour for a whole 2 min interval, points were awarded in this study contingent on good behaviour at the end of a 2 min interval. The differential effects of these two types of reinforcement schedule may be considered as a focus for future research, however the whole interval approach requires more teacher attention to award points appropriately. Groves and Austin also opted to take data only on individual students (both 'targets' and those considered well behaved). The type of data collected in the current study must be considered differently. In the current study, two target students were chosen based on teacher evaluation of their behaviour, however it was evident that the teacher could have chosen more target students from the class group, or could have chosen two different target students. The whole class data in the current study is therefore made up of well-behaved peers, but also students who may have engaged in similar levels of disruption to Adam and/or Ben.

Social Validity

The teacher in the current study found the CBGG acceptable and efficient, and rated it highly in these domains, however rated it slightly lower in effectiveness. She particularly liked how the game drew attention to children who were ‘always good’ and could sometimes go unnoticed, and she liked the team element. The teacher’s acceptability rating was 82 out of a possible 90. This is in line with other studies using similar measures of acceptability. Wright and McCurdy (2012) for example, had teachers fill out the IRP-15 (Martens et al., 1985), and the both teachers rated the CBGG 78 and 68 respectively. Ms. Daly rated the game substantially lower on effectiveness than for acceptability and efficiency (M = 3.29). Her effectiveness ratings appear to pertain to the generalisability of the intervention, as it is evident that she slightly disagreed with items which concerned the improvements in behaviour after the game was finished and in other settings. This is reflected in her comments, where she stated that when the game ended, the class would think they did not have to follow the rules anymore. This point is both a strength and a weakness. Its strength lies in the fact that it provides further evidence that the game and rules had clear control over the students’ behaviour, however, if this was a common outcome after the game has ended, it could be problematic for a teacher. This point has already been a focus of research by Donaldson et al., (2015), who demonstrated across five kindergarten classes, that the GBG during one activity did not produce sustained decreases in disruption in a proceeding activity. It has not yet been evaluated during the CBGG and this is a potentially useful avenue for future research. Future research should also potentially focus on investigating strategies teachers could put in place after the game has ended which may help to maintain appropriate student behaviour.

The BIRS (Elliott & Von Brock Treuting, 1991) was used to assess social validity by teachers in Chapter 3, Study 2. Conflicting ratings were given in Study 2, with Ms. Brady rating the CBGG negatively and Mr. Carroll rating the game positively, more similarly to Ms. Daly’s ratings in the current study. Mr. Carroll similarly provided the lowest ratings on the effectiveness scale (M = 4), and his ratings on the other two scales were very similar to those given by Ms. Daly. As noted in Chapter 3, Ms Brady’s ratings were very negative and based on her comments, seemed to be related to the time taken to implement the intervention. This concern was not apparent in the current chapter however it is important to appreciate the very different populations recruited across Chapter 3 Study 2 and the current study. The teachers in Chapter 3 taught adolescent students for short periods during the school day (i.e., 40 min) and implemented the CBGG during these short periods. Ms. Daly, being a primary school teacher, had her fourth-class students for the full school day. Therefore, the time taken to implement the intervention reflected a small percentage of the time which she taught the students, compared to the secondary school teachers. Ms. Brady also noted that some students actually appeared to behave worse during the game when observers were present. Data collection on individual students in her classroom may therefore have provided valuable insights. It should also be noted that there were differences in how the game was implemented across Chapter 3 and the current Chapter. For example, daily prizes were feasible in the current study and were not in Chapter 3. A different schedule was used for checking behaviour in the current study, with 2 min intervals between points compared to the 5 min intervals used in Chapter 3. The method used for setting a points criterion was also more objective in this study. Ms. Daly recorded prospective ‘points’ during one baseline session, which could then be used to set a realistic points criterion in intervention phases. This method was based on a similar approach taken by Ford (2017). This was a significant strength in this study compared to those reported in Chapter 3. These are all factors which could have influenced teacher ratings of the CBGG.

Limitations

Some limitations of the current study must be considered. First, just one class was recruited in one school setting, although efforts were made to recruit a second class to strengthen the research design. Efforts were made to recruit a single-sex female class for comparative purposes, however this wasn’t possible during the course of this study. A second limitation is that social validity measures were not obtained from the students themselves. It had been planned for these questionnaires to be administered at the end of the project, however the project ended abruptly when school closures were announced as part of the Irish government’s Covid-19 response. A third limitation lies in some aspects of the study not meeting WWC design standards fully (see Chapter 2 for a full description; WWC, 2017; 2020). IOA data were not collected in each phase for Ben as he was absent on the day a second observer was present in the second phase of the study. Nonetheless, IOA was still collected in over 20% of sessions overall for Ben. As well as this, only 3-4 data points were collected per phase in some cases, which means the study overall meets the WWC standards with reservations rather than meeting standards fully. The time frame during the day in which data were collected was only available 3-4 times per week, therefore, to finish the project in a timely manner, some phases needed to be shorter than five data points. It would have been desirable to collect data on teacher behaviour during this study, as had been done in Chapter 3 Study 2, to assess whether the CBGG impacted a primary teacher's rate of praise and/or reprimands (e.g., Tanol et al., 2010; Wahl et al., 2016). It was decided that these data would not be collected during this study as the game was to be played for 20 min and during this 20 min, data would need to be taken on the whole class as well as on two target students. Incorporating data collection on teacher behaviour within that small time frame proved unfeasible. It may be possible in a situation where data is not being taken on targets, or where the game is being played for a longer period. Finally, no data were taken to assess generalisation or maintenance effects. The latter was not possible as the school term ended abruptly. Generalisation data were not probed at different times during the school day, however, given the teacher's comment in her social validity ratings whereby she stated that students didn't follow the rules after the game had ended, it is possible that generalisation effects were not present.

Implications for Future Research and Practice

Research on the CBGG's effectiveness is clearly gathering pace, with recent demonstrations of its effectiveness in various populations, including with mainstream primary school children (Groves & Austin, 2020; Wahl et al., 2016; Wright & McCurdy, 2012), mainstream adolescent students (Ford, 2017; Chapter 3, Studies 1 & 2) and students with emotional and behavioural disorders (Groves & Austin, 2017). The current study further adds to the literature in support of the CBGG across whole class groups and with particularly disruptive students. It may therefore be considered more seriously by teachers for adoption as a positive classroom management intervention. This is particularly relevant to the single-sex boys' primary school setting in which this study was conducted, given the recent evidence specifying that punitive behaviour strategies are more commonly used in single-sex boys' primary schools than single-sex girls' schools or mixed-sex schools (Smith & Quail, 2017). The current study has further demonstrated the benefits of collecting data on individuals within the context of the whole class group. Previous research has demonstrated that sometimes individuals benefit more than the whole class during a group intervention (Hoff &

Ervin, 2013), but sometimes even if the whole class displays behavioural improvements, individual improvements may not be as large (e.g., Dadakhodjaeva et al., 2019). The latter was evident here to an extent, although individual students did demonstrate improvements overall.

One of the key points for future research which became apparent during this study, is the need for examination of fading strategies to maintain positive behaviour over longer periods and potentially when the CBGG is finished. Ms. Daly noted in her social validity ratings that when the CBGG ended, behaviour could deteriorate immediately. This has also been noted in previous research on the GBG, where student behaviour improved only during phases where the GBG was in place, and not during activities immediately following game cessation (e.g., Donaldson et al., 2015). Future research may examine the effects of reduced frequency implementation, which has been effective in the context of the GBG (Dadakhodjaeva et al., 2019). Fading the intensity of the intervention by thinning the schedule of reinforcement over the course of the game is another avenue for future research. This would mean that the game could remain in place, but less work would be involved for the teacher. This may allow the game to be played over longer periods of the school day and effects on behaviour could be prolonged. It was noted in the earlier Methods section of this chapter, that this study intended to probe schedule thinning during the CBGG by having the teacher implement the game with longer intervals between points. Unfortunately, due to government instructed school closures in March 2020, this was not possible. This is an important issue for future research and a systematic evaluation of schedule thinning during the CBGG is necessary. Although it has been discussed briefly before in the context of the CW-FIT, whereby the schedule of reinforcement was thinned over time during the group contingency component (e.g., Conklin et al., 2017), a systematic evaluation during the CBGG has not been conducted.

Conclusion

Overall, the current chapter has addressed a gap in the literature very relevant to Irish education contexts by targeting a single-sex boys’ school setting. Furthermore, collecting data on group and individual behaviour within this content provided rich data for evaluation. The CBGG was effective in increasing AEB and reducing DB with the fourth-class class group taking part here, as well as with two particularly disruptive target students, Adam and Ben. The study has expanded on previous research on the CBGG with middle primary school students (e.g., Wright & McCurdy, 2012) by examining these individual effects and comparing them with the larger class group. The study has pointed towards useful avenues for future research, including the examination of strategies for fading the game while maintaining effects on behaviour, and examination of the game in other single-sex settings.

Chapter 5: Thinning the Schedule of Reinforcement during the CBGG

Introduction

Positive reinforcement has been established as a common feature in classroom management practices, with it playing a role in praise interventions (Moore et al., 2019), token economies (Maggin et al., 2011) and of course game-based classroom management interventions, such as those at the focus of this thesis (see review in Chapter 2). Reinforcement is applied differently across classroom management interventions, with differing schedules of reinforcement applied. Some interventions also incorporate differential reinforcement, for example the CBGG (e.g., Wright & McCurdy, 2012). In understanding how behavioural classroom management intervention works, it is important to understand both schedules of reinforcement and differential reinforcement, and this Chapter will outline their application. In the previous chapter, it was apparent that perhaps the effects of the CBGG are not always sustained immediately after the game ends (according to teacher report). Manipulation of schedules of reinforcement during the game may provide an opportunity to fade the intensity of the game over time which may assist in maintaining effects while decreasing teacher workload.

Schedules of Reinforcement

A schedule of reinforcement is a rule describing the reinforcement contingency that determines the conditions by which behaviours will produce reinforcement (Cooper et al., 2007). Continuous reinforcement occurs when a response always produces reinforcement. For example, pressing the handle on an unlocked door will always be followed by that door opening. Intermittent reinforcement involves only some responses being reinforced, and there are different types of intermittent reinforcement schedule. These schedules may be response based (ratio schedules) or time based (interval schedules). As well as this, they may be fixed or variable. A fixed ratio (FR) schedule involves every *n*th response being reinforced (Ferster & Skinner, 1957). For example, a child may receive a token from a teacher after every 5 problems are completed (i.e., an FR5 schedule). A variable ratio (VR) schedule is like the FR schedule, however the amount of responses required to produce reinforcement is variable, unpredictable and averages around *n* (Cooper et al., 2007; Ferster & Skinner, 1957). For example, a VR5 schedule would see a child receive a token from a teacher after differing amounts of responses averaging around 5 (e.g., sometimes 5, sometimes 6, sometimes 1). Interval schedules of reinforcement are based around time. A fixed interval (FI) schedule provides reinforcement directly after the first response after a certain period of time has elapsed (Cooper et al., 2007). This period is fixed and begins directly after the preceding reinforcement (Ferster & Skinner, 1957). A variable interval (VI) schedule is very similar, however the time elapsed is not fixed and changes after each incident of reinforcement, averaging around *n* (Ferster & Skinner, 1957). Intermittent schedules of reinforcement are commonly incorporated into gamification as when implemented correctly, they can enhance player engagement (Linehan et al., 2015). In classroom management practices, intermittent reinforcement is more practical than continuous reinforcement for teachers given that a) they can't possibly see all responses by all students at one time and b) the time which would need to be dedicated to a continuous reinforcement schedule may be unrealistic in a classroom setting.

Just as reinforcement can be presented on a schedule, so can punishment. Punishment maintains maximum suppressive effectiveness on a continuous schedule (FR1; Cooper et al., 2007). Aiming to suppress disruptive behaviour using an FR1 schedule of punishment by reprimanding children every time they engage in an undesirable behaviour, a number of issues may present themselves. First, a teacher may inadvertently be reinforcing the undesirable response via reprimands, by giving attention to the misbehaving student. Second, a teacher is unlikely to observe every single incident of misbehaviour and therefore the punishment schedule is not as dense as they may think. Despite the fact that reinforcement is emphasised for use over punishment in school behaviour policies (National Educational Welfare Board, 2008) and in incidences of gamification (Linehan et al., 2015), it is still used in game-based classroom management interventions such as the GBG (fouls are administered on an FR1 schedule). However, student behaviour is reinforced if it occurs at a low rate (differential reinforcement of low rates of behaviour/DRL; Cooper et al., 2007). Differential reinforcement in the context of game-based interventions will be discussed next.

Differential Reinforcement

Differential reinforcement is a type of reinforcement schedule where one response class is reinforced and reinforcement for another response class is withheld (Cooper et al., 2007). There are several types of differential reinforcement; differential reinforcement of incompatible behaviour (DRI), differential reinforcement of alternative behaviour (DRA), differential reinforcement of low rates of behaviour (DRL) and differential reinforcement of other behaviour (DRO). DRO is a commonly used type of differential reinforcement in game-based classroom management literature (e.g., Groves & Austin, 2020; Wright & McCurdy, 2012). This involves delivering a reinforcer when a problem behaviour has not occurred during or at a specific time (Cooper et al., 2007). The former is termed interval DRO and the latter, momentary DRO. Both have been commonly incorporated in game-based classroom management research, which will be discussed in further detail next.

Differential Reinforcement During the GBG and CBGG

The GBG as it is traditionally applied, involves an FR1 schedule of punishment, such that every time a student breaks a rule, their team receives a foul. The issue discussed earlier however does apply, in that the teacher likely does not actually observe every single incident of misbehaviour which occurs during the class (Wright & McCurdy, 2012), thereby creating a VR schedule of punishment (Joslyn & Vollmer, 2020). In the CBGG, positive reinforcement is an important feature, in that students receive points contingent on appropriate behaviour or contingent on engaging in a behaviour *other than* disruption, at a specific moment in time, that is, DRO. It has been suggested that the type of reinforcement applied here is a fixed or variable momentary DRO schedule, depending on how exactly points are awarded (FM-DRO/VM-DRO; Wright & McCurdy, 2012). Other researchers have used a fixed whole interval DRO schedule (Groves & Austin, 2017; 2020). This is a schedule whereby reinforcement is contingent on the absence of problem behaviour during a whole preceding interval. Both versions have been effective during the CBGG. Chapters 3 and 4 in this thesis have used a momentary DRO with decreases in DB and increases in AEB observed while the CBGG was in place. In Chapter 3, Studies 1 and 2 used a VI5-momentary DRO and the study in Chapter 4 used an FI-2 momentary DRO. Groves and Austin (2017; 2020) demonstrated that the game was effective using a FI2 whole interval DRO. In the 2017 study, several types of disruption were reduced across four individual students in a Pupil Referral Unit for children with emotional and behavioural disorders. They had been excluded from mainstream education due to their high rates of disruptive behaviour. The CBGG was trialled here via an interdependent group contingency (as it is usually played) and as an independent contingency (i.e., the students received points for themselves contingent on their own behaviour). Both were effective, although most of the students preferred the interdependent version. Groves and Austin (2020) demonstrated the efficacy of the CBGG with a FI2 whole interval DRO with six students in a mainstream Year 4 class. In other studies, it is unclear whether a schedule of reinforcement was put in place by the researchers, or whether it was simply up to the teacher themselves as to when they would award points (e.g., Lynne et al., 2017; Tanol et al., 2010).

Previous studies on the CBGG have set schedules of reinforcement for the duration of the study without changing the schedule over time. For example, Wright and McCurdy (2012) used a VI4 momentary DRO schedule for behaviour checks and this was not changed/lengthened throughout the CBGG phases despite behaviour improving across the class. Other studies have used even denser schedules. For example, when examining the efficacy of the CBGG with a high school class, Ford (2017) had the teacher prompted every 2 min to scan the classroom. The teacher was encouraged to scan the room between 0-30 seconds after the notification to ensure there was some variability in the schedule, however the schedule remained dense throughout the study. In the Quiet Classroom Game, a game similar to the CBGG which was identified in the review in Chapter 2, the teachers awarded points to their classes contingent on low noise levels at the end of 2 min intervals (Radley et al., 2016). This is an interval DRL schedule (i.e., reinforcement is contingent on *low levels* of noise rather than absence of noise) rather than a DRO schedule, but similar principles apply in that reinforcement is in place at the end of an interval and the reinforcement schedule is relatively dense. In the current thesis, different schedules of reinforcement have been used during CBGG implementation. In Chapter 3, there were 5 min intervals between opportunities to earn points, and in Chapter 4, intervals were 2 min in length. These decisions were made in conjunction with class teachers and took the class context into account. That is, 2 min intervals were feasible for use with a fourth-class group as the teacher taught the same group all day, but the second-level teachers only had up to 40 min with their class groups, making 5 min a more realistic interval length.

Schedule Thinning

Schedule thinning has been applied systematically in the past to fade out interventions and to make them less time intensive for the intervention agent. Benefits of decreasing intervention intensity include enhancing the ease of management of reinforcement delivery for the intervention agent and making the intervention more naturalistic and therefore potentially more generalisable (LeBlanc et al., 2002). For example, in a novel application, Bergstrom et al. (2011), demonstrated the efficacy of DRO in reducing inappropriate touching of a pet dog by a 6-year-old boy with autism spectrum disorder. Initially, the boy received a chosen reinforcer at the end of a 10 s interval if he had not touched the dog inappropriately during the preceding interval. This interval was increased to 15, 20, 30, and eventually 600 s (10 min). The behaviour was effectively eliminated through the thinning of the reinforcement schedule over time. Toussaint and Tiger (2012) examined the efficacy of a VM-DRO with a 12-year-old boy who engaged in covert self-injurious behaviour in the form of skin picking. The DRO schedule was initially set around an average of 8 s. The boy would be left alone in a room and a therapist would enter and deliver praise and a token contingent on the absence of skin picking at the exact moment of entry. The schedule was gradually lengthened with the mean interval increasing by 50% after two consecutive sessions where skin picking occurred at a rate 90% below baseline. Following a brief withdrawal phase, the DRO eventually was faded to 300 s, maintaining skin picking at a very low to near zero rate.

In the classroom management literature, particularly that around the popular GBG/CBGG, interventions are often put in place without allowing for a period of schedule thinning or intervention fading. This means that teachers or other intervention agents are potentially implementing an intervention at a level much more intense than needed, which depletes valuable resources. One exception was a study by Conklin et al., (2017), which evaluated the CW-FIT, a package intervention which incorporates a group contingency game very similar to the CBGG. In this study, the authors did incorporate a schedule thinning procedure whereby the teacher began by awarding points to teams every 1-2 min and thinned the schedule to every 3-5 min over time as students became more proficient at prosocial skills taught as part of the game. The authors did not go into detail on the schedule thinning procedures, nor did they specify at what point schedule thinning occurred so that effectiveness of the procedure can be analysed. Naylor et al., (2018) conducted a similar study on the CW-FIT, in which they state schedule thinning was incorporated into the group contingency game procedure. Similarly to Conklin et al., (2017), they did not outline at what stage they began to thin the schedule, simply that they began with a dense schedule (1-2 min) and progressed to a leaner schedule (1-4 min). It is clear that more explicit descriptions of schedule thinning procedures are needed in the game-based classroom management literature to aid with replication.

In Chapter 3 in the current thesis, 5 min intervals were chosen on the basis that they were the most realistic interval length for the teachers to employ during a 35-40 min class with adolescent students. In Chapter 4, 2 min intervals were used as the setting allowed for a denser schedule. No study on the CBGG has explicitly examined the effect of changing the schedule of reinforcement over time despite recent calls for evaluations of procedural variations of the game (Joslyn et al., 2019). An application of a schedule thinning procedure during the CBGG has the potential to occupy a distinct and important gap in the literature. Published studies on similar game-based interventions have changed the schedule over time (e.g., Conklin et al., 2017; Naylor et al., 2018), however detailed accounts of the timing of the schedule changes are not provided. Providing a detailed account of the steps taken during the schedule thinning procedure is important to inform practice, as behaviour analytic applications should be technological, that is, the techniques applied in a study should be identified and described completely (Baer et al., 1968). There is therefore scope to examine the CBGG with younger populations, beginning with a dense schedule and progressing to a thinner schedule over time.

Research Questions

A number of research questions are proposed:

1. Is the CBGG, with 2 min intervals between points, effective in targeting academically engaged behaviour and disruptive behaviour in a mixed-sex, mainstream senior infants classroom environment?
2. Is the CBGG, with 2 min intervals between points, effective in targeting academically engaged behaviour and disruptive behaviour in two teacher-selected students displaying high rates of disruptive behaviour and/or low levels of engagement?
3. Is the CBGG's effectiveness maintained across the whole class and individual target students when the reinforcement schedule is thinned from 2 to 3, 4 and 5 min?
4. Does the teacher find the CBGG an acceptable and effective intervention for application with a mixed, mainstream senior infants class?
5. Do part-taking students find the CBGG to be an acceptable intervention?

Method

Recruitment, Participants and Setting

Recruitment for the current study was identical to that in Chapter 4, in that the school and teacher identified for participation were contacted via the same email procedure. Similarly to Chapter 4, a principal made contact from one of the emailed schools, noting that a senior infants teacher was interested in taking part in the study. A meeting was organised with this teacher where the project was outlined and she was given relevant recruitment materials, identical to those used in Chapter 4. The teacher was then given recruitment packs to send home with the children in her class containing PLS's and consent/assent forms. The parental consent forms were identical to those used in Chapter 4 and the student assent forms were modified slightly to suit the young, senior infants population (Appendix Y). As in previous studies discussed, only students returning completed parental consent and assent forms would participate in having their behaviour monitored, but the teacher would implement the classroom management intervention with the whole class.

The school in which this study took places was a single-sex girls' primary school in an urban area. The part-taking teacher, Ms. Ellis (pseudonym), was a 48-year-old female with 12 years of teaching experience. The senior infant class consisted of 26 consenting students (15 female, 11 male). Although the school was a single-sex girls' primary school, junior and senior infant classes were mixed-sex. After these initial two years of schooling are completed, the boys move to an all-boys school setting. Senior infants is the second year of formal schooling in the Irish education system and is approximately equivalent to kindergarten in the US school system or Year 1 in the English school system. Students in senior infants are approximately aged 5-6 years. Senior infants attend school for one hour less per day than older primary school students (i.e., first class and older). Ms. Ellis chose two individual students for individual behaviour monitoring based on her evaluation of class behaviour. Target students were also chosen based on their attendance record in that the teacher was asked to choose students who would usually be present during data collection sessions. Ellie was a five-year-old female student and Katie was a six-year-old female student (pseudonyms).

Materials

The teacher was provided with a laminated copy of the class rules (Appendix Z) and five scoreboards, one for each table in the classroom. The scoreboards were coloured penguins with spaces for points to be coloured on when administered (Appendix AA). Each penguin was coloured to correspond with a table, as each of the five tables in the room had an assigned colour. The teacher was given an Octopus watch which was pre-programmed to vibrate and remind her to scan the classroom intermittently. The teacher chose the prize based on what she had used with students before. It was also important that the prize was different to anything being used by the teacher in other classroom management procedures throughout the day. Prizes used were stamps and stickers, which were cost effective and appeared desirable to the students. The teacher stuck a checklist to the wall which outlined the steps of the game. This checklist was identical to that used in Chapter 4. Observers used the same checklist to carry out treatment integrity checks. Data was collected by observers with paper and pen and interval changes were signalled via earphones.

Operational Definitions of Target Behaviours/Dependent Variables

Data were collected on academically engaged behaviour (AEB) and disruptive behaviour (DB) in this study. Both are described in more detail here.

Academically Engaged Behaviour (AEB)

> *"A student is academically engaged if they are giving their attention to the academic task which is ongoing. This includes writing, colouring, reading aloud or to oneself, conversing with a peer about the task (where this has been permitted by the teacher), eye contact is oriented towards the task or teacher or the student is using the class sharpener or walking in the direction of the sharpener at the time of recording. A student is not considered engaged if they are walking to or from the sharpener and are engaging with a student at another table at the time of recording. A student is also not considered engaged if an out-of-seat behaviour, verbal disruption, or other motor disruption is ongoing at the time of recording. Exceptions to this is when a student is engaging in motor disruption which is compatible with the academic task e.g. the student is playing with a pencil, but simultaneously looking at the board".*

Disruptive Behaviour (DB)

DB was monitored across three different categories, and these categories were identical to those used in Chapter 4: verbal disruption (VD), out-of-seat behaviour (OOS) and motor disruption (MD). Definitions of each behaviour differ slightly to those used in Chapter 4 as behaviour manifested slightly differently in the younger age group taking part here. As in Chapter 4, when analysing data each category of DB is considered separately as well as the composite variable of DB. Each category of disruption is defined here.

Verbal Disruption.

> *"A student is considered to be engaging in VD if they engage in a vocalisation not authorised by the teacher and unrelated to the work ongoing in the classroom, for example singing, whistling, humming, talking to a peer about something unrelated to the work, shouting out when a question has not been posed by the teacher (i.e., sometimes the teacher will pose a question or statement and allow/expect students to*

respond by shouting out an answer or engaging in choral responding), or shouting out something unrelated to the academic task".

Out-of-Seat Behaviour.

"Out-of-seat behaviour is defined as a student leaving a seated position during classwork and moving > 1 metre from their chair. Exceptions include instances where students leave their seat to walk directly to the pencil sharpener/pencils and directly back again and instances where students have raised their hand, asked for and received permission to leave their seat from the teacher. If a student leaves their seat to sharpen their pencil and engages with a student sitting in a different group on the way/on the way back, this is marked as an incident of OOS. The type of behaviour engaged in while out of seat is also marked".

Motor Disruption.

"Motor disruption involves playing with objects (i.e., using one or two hands to manipulate an object) in a manner which is incompatible with the academic task, turning in one's chair away from the task (>3 seconds), leaving one's head on the desk, swinging on two legs of the chair or physically interacting with a peer in a manner which is incompatible with the academic task".

Observation Procedures and Data Collection

Data were collected up to four times per week in the classroom during the last period of the school day. This was from immediately after lunch break at 1.10pm, until the students started to get ready to go home at 1.40pm (home time was at 1.50pm). Data collection sessions lasted 20 min and during this period, the class engaged in a mathematics lesson. The session involved the teacher explaining a concept on the board for 4-5 min followed by independent seatwork for 15-20 min. Students could converse appropriately during this independent seatwork because helping each other out was encouraged. As with all previously outlined studies in this thesis, data was collected via momentary time sampling and partial interval recording, and an individual fixed method was used. Similarly to Chapter 4, intervals were set to 10 s, with 5 s to record. Data collection was identical to Chapter 4, in that observers alternated between observing a student in the class, followed by a target student every second interval. Data were collated in a manner identical to Chapter 4.

Observer Training and Interobserver Agreement

Second observers were trained for the studies outlined in Chapter 4 and 5 simultaneously, therefore observer training was identical to the outlined training in Chapter 4. IOA were collected on 38.46% of observation sessions overall and during at least 20% of baseline and intervention phases, as per WWC recommendations (WWC, 2017; 2020). It was collected at least once per phase for each outcome. Ellie was absent during two sessions where IOA data were collected and therefore IOA data was not collected for her during one of the intervention phases (the B4min phase). Katie was absent during the very last data collection session, therefore IOA data were not collected for her behaviour during this phase (the B5min phase). Therefore, IOA was collected on 38.46% of occasions for the whole class (10 out of 26 sessions), 33.33% of occasions for Ellie (8 out of 24 sessions) and 37.5% of occasions for Katie (9 out of 24 sessions). IOA for student behaviour was calculated using interval-by-interval agreement and dividing the number of agreements by the total number of observation sessions and multiplying by 100 to obtain a percentage. Table 5.1. outlines the average IOA and range for each participant and outcome.

Table 5.1.

Mean IOA and Range Across Participants and Phases

Outcome	Whole Class *M* (range)	Ellie *M* (range)	Katie *M* (range)
Verbal Disruption	91.53% (81.58-97.5%)	86.64% (75-100%)	88.36% (75-100%)
Out-of-Seat	95.83% (90.63-100%)	94.82% (75-100%)	96.04% (86.67-100%)
Motor Disruption	90.91% (82.5-96.88%)	90.08% (85-100%)	84.94% (65-100%)
Total Disruption	92.75% (88.33-95%)	90.34% (85-94.12%)	89.67% (80-97.92%)
Academically Engaged Behaviour	87.38% (70.27-97.5%)	84.55% (80-90%)	83.06% (75.32-95.46%)

Social Validity

Following the final day of data collection, the teacher and students completed social validity measures. Ms. Ellis completed the BIRS (Elliott & Von Brock Treuting, 1991) and students completed a modified version of the CIRP (Mitchell et al., 2015; Witt & Elliott, 1985). The BIRS was identical to the version used in Chapter 2, Study 2 and in Chapter 4, with some additional questions regarding the schedule thinning procedure. Ms. Ellis was asked to answer the BIRS questions with reference to the CBGG generally. The additional questions were as follows; "Did you have a preference for any particular version of the CBGG (2, 3, 4 or 5 minutes between points)?"; "Did you think any particular version was more/less effective than another?"; "Did you think any particular version was easier/more desirable to implement than others?"; "Do you have any further comments/feedback on the CBGG?".

The modified CIRP is a social validity measure with 8 items such as "Did you like participating in the Game?", to which students answered 'yes' or 'no'. The measure was similar to that used in all of the previous studies outlined in this thesis, however more substantial modifications were made for use in this study to make it suitable for the senior infants population (Appendix BB). This included the rephrasing of some questions and including smiley faces and thumbs up/down as options to indicate approval/disapproval. As in previous studies, the highest rating a student could give the game was eight. If a student responded negatively to a question, the researcher would ask them why and write down their responses rather than having them write out their own comments. This was considered appropriate given the age of the population.

Experimental Design

A reversal/withdrawal design with phases A-B2min-A-B2min-B3min-B4min-B5min was used to evaluate the CBGG across the class and individual students taking part. The A phase refers to baseline, B2min refers to the CBGG with 2 min intervals between opportunities to earn a point, B3min refers to the CBGG with 3 min intervals between opportunities to earn a point, B4min refers to the CBGG with 4 min intervals between opportunities to earn a point and B5min refers to the CBGG with 5 min intervals between opportunities to earn a point. Phases B3min-B5min were considered an extension to the B2min phase as nothing else changed except for the interval length between teacher behaviour checks, that is, opportunities to earn a point.

Procedure

Baseline

During baseline, the teacher proceeded with planned educational tasks with no intervention in place. There was no specific contingency in place for rewarding positive behaviour during educational tasks and disruptive behaviour was addressed with verbal warnings. The teacher used a positive reinforcement contingency for transitions at the beginning and end of instructional periods. A countdown timer was put on the board and students had to be ready at their tables by the time the timer reached zero. The teacher then awarded a star to the ‘best table’. This continued throughout the day every day of the week and the table with the most stars on a Friday afternoon could choose a toy to play with from a toy box. This strategy was not considered to interfere with the implementation of the CBGG, so the teacher continued to do this throughout all phases. Like in Chapter 4, during the last day of baseline data collection, the teacher was asked to privately ‘play the game’, without letting the children know about it to allow her to get used to the watch, practice the game procedures privately and to assist with setting a criterion for intervention phases.

Teacher Training

Teacher training was identical to that conducted in Chapter 4 and took place during one session after school the day before baseline data collection was finished. Ms. Ellis was shown the behavioural data for her class and for the individual targets within the class. All other elements of the training were the same, however a different reward was chosen by this teacher. A preference assessment was due to be carried out, however the teacher already used many of the potential reinforcers in her transition classroom management strategy. She therefore decided there were not enough alternatives which could be put on a preference assessment questionnaire and that stamps and stickers would be an appropriate and cost-effective reward for the group. Ms. Ellis was introduced to the concept of schedule thinning. She was trained in implementation of the CBGG with 2 min intervals between behaviour checks. She was told that short meetings introducing her to thinned schedules would take place before each schedule change.

Intervention: Caught Being Good Game

The CBGG was introduced to the class during the first session following completion of baseline data collection. The class was divided into five teams which were colour coded based on their table, that is, the Red team, the Blue team etc. Each table served as a separate team. The teacher told the children about the game, introducing it as the 'Penguin game', to tie in with the fact that the scoreboards were created with a penguin theme. This aligned with ongoing work on Antarctica in the class. Students were told that during the game there would be ten chances to earn a point for following the class rules. These rules were as follows:

- Look at and Listen to your teacher
- Hands up and wait for your teacher
- Do your best at your work
- Respect your friends & let them do their work
- Stay in your seat

Like in Chapter 4, the game would take place for 20 min (one point every 2 min for 20 min = 10 points available in total).

Ms. Ellis announced that the game had started after explaining how the game was played and reviewing the class rules. When the teacher was prompted to scan the room for behaviour and award points if appropriate, she would do so on the corresponding penguin scoreboard for each team, by colouring in one of ten buttons on the penguins' torso. The goal for the 2 min version of the game was 7 points. This was calculated in the same way as in Chapter 4, by taking the average points earned by each team during the last day of baseline data collection and adding 10%. The average amount of points earned during the last day of the baseline phase was 6.2. Adding 10% to this gave a total of 6.82 and this was rounded to a goal of 7 points, or 70% of the total points available.

Prizes were awarded daily immediately after the game. Teams meeting or exceeding the 7-point goal were eligible to receive the prize. The prize was a choice between one of two stamps, or a sticker. The stamps were brightly coloured markers which would make different shaped and coloured stamps, for example, a red apple, an orange star. The choice given daily was varied in that the teacher chose two different markers from a set of ten. The team with the most points got to choose their stamp or sticker first.

Schedule Thinning. After the initial ABAB phases where the game was implemented with 2 min intervals between points, the schedule of reinforcement was thinned over time. Guidance was taken from LeBlanc et al., (2002) in planning these phases. The intervention was applied at an intensive level to begin with, with 2 min between reinforcement opportunities. One min was considered but was deemed unrealistic for a teacher of a large class of mainstream students to maintain. The method of schedule thinning was decided upon, in that it was decided to increase the time between behaviour checks in 1 min increments. Although this is a large percentage change in the schedule (an increase of 50% initially from 2-3 min, 33.33% increase from 3-4 min and 25% increase from 4-5 min), it is still a short time frame and was deemed easy to understand and to remember for the teacher implementing the intervention. The terminal goal would be 5 min. LeBlanc et al., (2002) recommend two consecutive exposures to a new schedule before proceeding to the next level. To align with this, and to align with the WWC recommendations of at least 3 data points per study phase (WWC, 2017; 2020), it was planned to expose the students to a new schedule over three consecutive days before proceeding to the next level. Table 5.2. summarises the phases, maximum points available in each phase and points criterion for each phase. It was planned that the game would be implemented with 5 min intervals for one more day, however the school term ended abruptly, and data collection ceased early, due to school closures amid the Covid-19 pandemic.

Table 5.2.

Description of Intervention Phases Across the Study

Phase[a]	Length of Schedule	Maximum Points Available	Criterion
B1	2 min	10	7
B2	2min	10	7
B3min	3 min	7	5
B4min	4 min	5	3
B5min	5 min	4	3

[a] B phases refer to intervention phases. B1 and B2 refer to the first two intervention phases, with 2 min intervals between opportunities for points. The subsequent B phases have longer interval lengths as outlined in the table.

Treatment Integrity

The teacher checklist for game implementation was almost identical to that used in Chapter 3, Study 2 and in Chapter 4. Ms. Ellis stuck the checklist to her classroom wall to serve as a prompt to complete each step. Treatment integrity data was collected during 100%

of intervention data collection sessions by the student researcher. If treatment integrity dropped below 80% for more than one day in a row, this was brought to the teacher's attention via email or in person and she was encouraged to use her checklist and complete each step. Overall mean treatment integrity across all intervention phases was 76.14% (range = 45.45-100%). Mean and range of treatment integrity by phase are presented in Table 5.3.

Table 5.3.

Mean Rates and Range of Treatment Integrity during Intervention Phases for Ms. Ellis

Phase	*M*	Range
CBGG 2 min 1	79.55%	54.55-100%
CBGG 2 min 2	74.55%	54.55-90.91%
CBGG 3 min	87.88%	81.82-90.91%
CBGG 4 min	78.79%	72.73-81.82%
CBGG 5 min	63.64%	45.45-81.82%

The most commonly missed step by the teacher was Step 2, that was, reminding students of the team divisions. This was not deemed hugely problematic however, as the teams were based on the students' tables which were colour coded, as were the corresponding score boards. The students also engaged in many other activities as a table group, so it was likely that students understood which team they were on without a reminder. The teacher also often forgot to review the game rules (Step 4). This was only completed during 47.06% of intervention sessions. Other steps completed less than 70% of the time included reminding students of how many points they needed to get a prize and explicitly stating how points were earned. All other steps were completed during more than 80% of intervention sessions.

Data Analysis

Data were analysed in the same way as previous studies, taking guidance from the WWC design and evidence standards (WWC, 2017; 2020). Effect size was calculated using Tarlow's recommendations for calculation of Tau (Tarlow, 2017), as was conducted in all previous studies.

Ethical Considerations

The ethical considerations and protocols in this study were identical to those considered in Chapter 4. No further ethical amendments were required to conduct this study. Modified student assent forms were used to ensure they were suitable for a senior infants population and some study materials were modified (e.g., the CIRP) to suit this population of young children.

Results

WWC Design Standards

The study design was in line with the WWC design standards as far as possible (WWC, 2017; 2020). The CBGG (i.e., the independent variable) was systematically manipulated throughout the study. IOA data were collected at least 20% of the time overall and in each condition, and at least once per phase for each outcome. The only exception to this was in the final B5min phase, where Katie was absent during the observation session when the second observer was present. This was not considered hugely problematic as IOA data for Katie was sufficient overall. IOA was above the 80% threshold for all participants and outcomes. There were at least three attempts to demonstrate intervention effects. Some phases contain less than five data points, and therefore the study meets design standards with reservations.

Visual Analyses

Class-wide Effects

Data on class-wide AEB, DB and sub-types of DB across phases are presented in Figures 5.1. and 5.2.

Academically Engaged Behaviour. During the first baseline phase, AEB occurred during a mean of 66.46% of intervals (range = 53.33-80%). AEB was highly variable during this phase. When the CBGG was introduced with 2 min intervals, AEB increased immediately and stabilised across the phase (*M* = 88.78%, range = 84.85-92.5%). There was no overlap between this intervention phase and the preceding baseline phase.

The CBGG was withdrawn in the following phase and AEB decreased immediately and substantially (*M* = 67.76%, range = 62.5-76.32%). Again, there was no overlap between this phase and the intervention phase which preceded it. AEB remained low and quite stable across the phase with no obvious trend in the data. When the CBGG was reinstated, AEB increased substantially (*M* = 87.79%, range= 81.58-97.37%). There was no overlap with the previous withdrawal phase.

During the B3min phase, AEB remained high (*M* = 88.52%, range = 80.56-97.5%), particularly across the first two data points. There was a decrease during the third data point in this phase, however AEB did not decrease enough to overlap with either of the baseline phases. During the B4min phase, AEB remained high and stable (*M* = 86.79%, range = 82.5-92.5%) and occurred at a similar level to previous intervention phases. Finally, during the B5min phase, the CBGG was implemented across two sessions with 5 min intervals between points. AEB remained high during the first data point in this phase, however decreased during the second data point (*M* = 84.59%, range = 79.17-90%). Treatment integrity was low during the second data point in the B5min phase (45.45%) and it was the only intervention data point across all phases to overlap with any baseline data point. There was no opportunity to collect a third data point for this phase.

Disruptive Behaviour. Three separate categories of disruptive behaviour will be discussed here: VD, OOS and MD (see Figure 5.2.), followed by a discussion of DB as a composite variable.

Verbal Disruption. VD occurred at a high and increasing rate during the first three data points in the initial baseline phase, before decreasing during the last data point (*M* = 25.83%, range = 12.5- 43.33%). When the CBGG was implemented, this behaviour reduced further and remained low and stable (*M* = 4.91%, range = 2.78-9.09%). During the withdrawal phase, VD increased immediately, and a steady increasing trend was observed across much of the phase (*M* = 15.63%, range = 10-25%). Verbal disruption decreased when the CBGG was reintroduced and remained low across the phase (*M* = 8.76%, range= 2.63- 23.68%), with the exception of the third data point where this behaviour increased substantially during one session (data point 16; see Figure 5.2.). When the schedule thinning phases commenced, VD remained low and stable across all three phases; the B3min (*M* = 1.67%, range = 0-2.5%), B4min (*M* = 3.31%, range = 2.44-5%) and B5min phases (*M* = 2.09%, range = 0-4.17%).

Out-of-seat Behaviour. OOS was not a huge issue with this class, and at baseline, it occurred during an average of 7.92% of intervals (range= 2.5-16.67%). OOS decreased when the intervention was introduced and remained low and stable across the first CBGG phase (*M*= 2.15%, range = 0-3.03%). When the intervention was withdrawn, there was a slight increase in

OOS, however there were two data points in this phase where OOS did not occur at all (*M* = 3%, range = 0-7.5%). When the CBGG was reintroduced, OOS remained low and stable (*M* = 1.03%, range = 0-2.63%) and this pattern was evident across the three subsequent phases where schedule thinning was introduced; B3min (*M* = 1.76%, range = 0-2.78%), B4min (*M* = 0.81%, range = 0-2.44%) and B5min (*M* = 2.09%, range = 0-4.17%).

Motor Disruption. MD occurred at a relatively stable rate during baseline, with one data point much higher than the other three (*M* = 17.92%, range = 10-36.67%). When the CBGG was put in place initially, MD decreased (*M* = 7.7%, range = 2.5-16.67%) and remained stable across the first three data points in the phase. There was an increase during the last data point in this phase. When the CBGG was withdrawn, MD occurred at a similar level to the final data point in the previous phase and increased further during the last two data points in the phase (*M* = 17.66%, range = 12.5-25%). When the CBGG was reintroduced, MD decreased immediately and occurred at a stable rate across the phase (*M* = 9.13%, range = 2.63- 12.5%). MD remained low during the B3min phase (*M* = 5.19%, range = 0-10%). This rate was slightly higher during the B4min phase, however not as high as during baseline phases (*M* = 11.59%, range = 9.76-15%). Finally, in the B5min phase, MD occurred during a mean of 7.5% of intervals (range = 6.67-8.33%). This was a low rate when compared with baseline phases.

Disruptive Behaviour Composite Variable. At baseline, class-wide DB occurred at a high rate (*M* = 34.38%, range = 27.5- 50%). The behaviour was quite stable, with one data point noticeably exceeding the other three (data point 3; see Figure 5.1.). DB reduced immediately upon introduction of the CBGG (*M* = 12.61%, range = 7.5-16.67%). This reflects a large reduction and there was no overlap here with the initial baseline phase.

There was an immediate and moderate increase in DB when the CBGG was withdrawn (*M* = 27.71%, range = 21.05- 37.5%). There was steady increase across most of the phase and no overlap with the previous intervention phase. There was an immediate decrease in DB when the CBGG was reintroduced (*M* = 15.32%, range = 2.63- 28.95%). There was one data point here where DB occurred at a rate similar to baseline phases (data point 16; see Figure 5.1.). It is evident that a high rate of verbal disruption took place during that observation. DB remained low and stable when the schedule was thinned to 3 min between opportunities to earn points (*M* = 7.78%, range = 2.5-12.5%). This reduction was sustained throughout the ensuing B4min phase (*M* = 11.59%, range = 9.76-15%) and the B5min phase (*M* = 11.67%, range = 6.67-16.67%). There was no overlap in phases B3min, B4min or B5min with either of the baseline phases.

Figure 5.1.

Percentage of Intervals with AEB and DB Across Study Phases for Ms. Ellis' Class Group

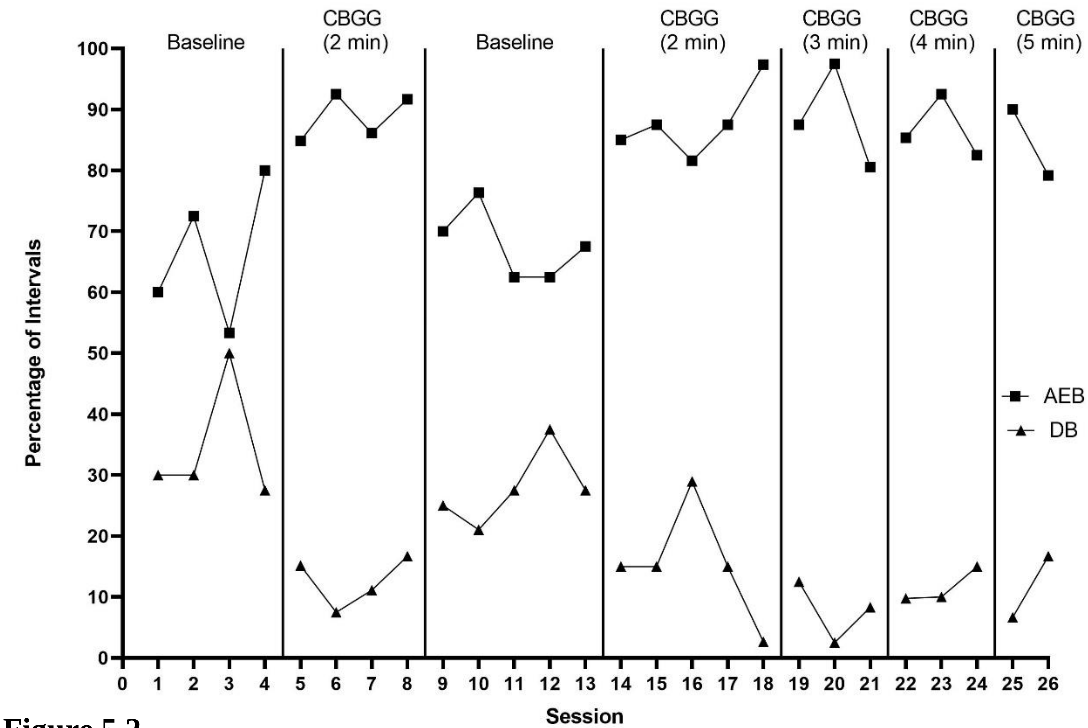
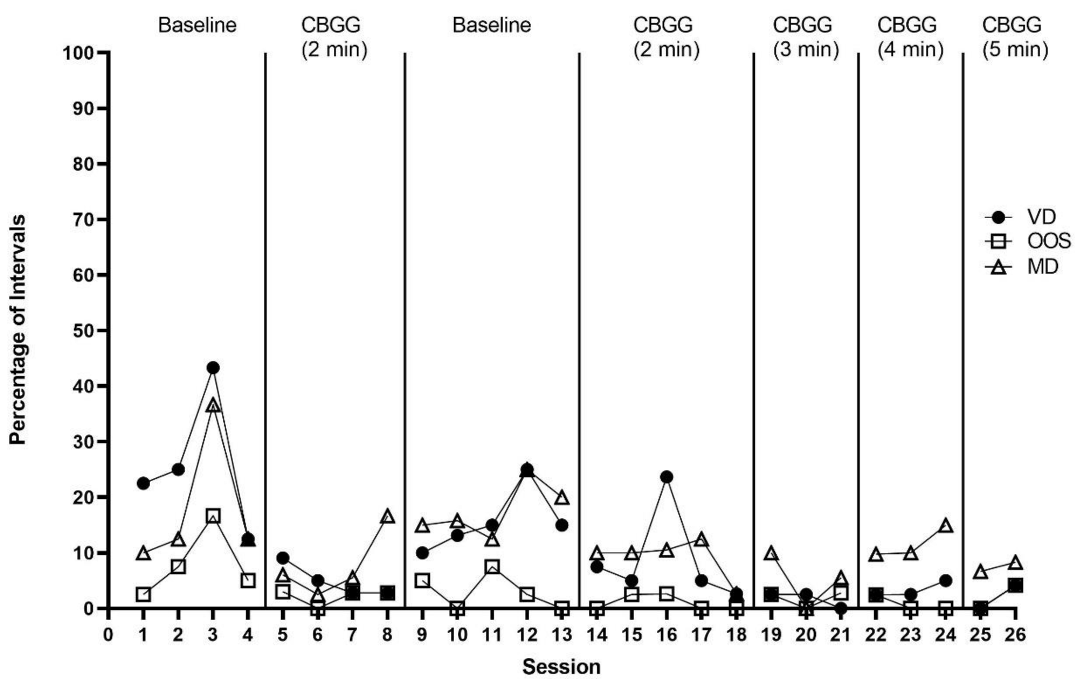


Figure 5.2.

Percentage of Intervals with types of DB Across Study Phases for Ms. Ellis' Class Group



Target Students

Data on AEB, DB and sub-types of DB for Ellie and Katie are presented across Figures 5.3. to 5.6.

Academically Engaged Behaviour.

Ellie's rate of AEB at baseline followed a similar pattern to the whole class in that it was variable and generally low (*M* = 66.25%, range = 60-80%). When the CBGG was introduced, Ellie's AEB increased immediately and substantially and was sustained across three data points. AEB decreased again towards the end of the phase (*M* = 82.6%, range = 72.22- 94.44%), in contrast to the whole class AEB which remained high. The final data point in this phase overlapped with one data point in the baseline phase.

During the first withdrawal phase, Ellie's AEB was highly variable and was not similar to the class, but despite this variability, her AEB was higher than the whole class on average (*M* = 75.33%, range = 60-85%). Ellie's rate of AEB was high during the first data point and during the last two data points. The overall level of AEB remained lower than during the previous intervention phase, however there was a high degree of overlap. When the CBGG was reintroduced, there was an increase in AEB (*M* = 92.44%, range = 85-95%). Again, her level of AEB was higher than that of the class and just one of the four data points in this phase overlapped with the preceding baseline phase.

During B3min, Ellie's AEB occurred during a mean of 86.3% of intervals (range = 80-90%), which demonstrates a slight decrease in level when compared to the previous B2min phase. Behaviour was similar to the whole class here. During the B4min phase, AEB was high across the two sessions for which Ellie was present (*M* = 92.5%, range = 90-95%). In the final B5min phase, AEB was variable, occurring during a mean of 82.08% of intervals (range = 70.83- 93.33%). The low rate of AEB during the final data point corresponds with low AEB across the class and low treatment integrity by the teacher.

Katie's rate of AEB was variable across the initial baseline phase, with a decreasing trend across the final three data points (*M* = 55.83%, range = 35- 80%). Her mean level of AEB was much lower than that of the whole class. An immediate and large increase in AEB occurred when the CBGG was introduced (*M* = 92.36%, range = 88.89-94.44%), such that Katie's AEB was higher than the whole class mean for this phase. It must be noted here that at the beginning of this CBGG phase, Katie moved seats to the other side of her table. When the game was introduced, this move was requested by her classmate who she spoke to a lot (at another close by table). Katie agreed to it, agreeing that it would help her earn more points. Ms. Ellis allowed this seat change upon the students' request. Although this change of seats serves as a confounding variable, it came about as a direct result of the game being introduced, at request of the students. During the rest of this phase, when the CBGG was introduced, Katie would move her chair to the other side of her table without asking.

When the CBGG was withdrawn, a request was put to Ms. Ellis by the researcher to keep Katie in her 'CBGG seat' during observations to see if behaviour would revert to levels similar to the initial baseline phase. This was deemed important to differentiate between improvements in behaviour due to the seat change versus the CBGG. Ms. Ellis agreed to this request. During this withdrawal phase, Katie's AEB immediately decreased and remained low across the first three data points before increasing towards the end of the phase (*M* = 74.22%, range = 61.11- 90%). During the fourth data point in this phase, a new seating plan was put in place for all students and everyone was moved to a different seat. Katie's increase in AEB during this phase coincided with the seating plan change. When the CBGG was reintroduced, Katie's AEB was initially similar to the withdrawal phase, before decreasing substantially for two data points, and increasing again during the final data point (*M* = 84.18%, range = 72.5- 94.74%). Despite the instability, the overall level of AEB was higher than during the preceding withdrawal phase and the overall level was similar to that of the whole class.

The schedule thinning phases saw Katie's AEB remain relatively high and stable. During the B3min phase, AEB occurred at a mean rate of 92.96% (range = 88.89- 95%). The seating plan changed again at the beginning of the B4min phase. During this phase, AEB remained high and occurred at a mean rate of 91.54% (range = 84.62-95%). Katie was only present for one of the two observations during the B5min phase, and AEB occurred during 73.33% of intervals. This reflects a moderate decrease when compared with the B4min phase.

Disruptive Behaviour. Target students' DB will be discussed here under three distinct headings (VD, OOS, MD; see Figures 5.5. and 5.6.) before being discussed as a composite variable (see Figures 5.3. and 5.4.).

Verbal Disruption. Ellie's engagement in VD was variable at baseline (*M* = 25.83%, range = 15-35%) with low and very high rates reflected across the phase. When the game was introduced, her rate of VD remained relatively high, however there was an overall drop in level (*M* = 17.88%, range = 5.56-25%). During the withdrawal phase, her VD initially decreased, before sharply increasing across the phase (*M* = 22.44%, range = 5-50%). Again, the data were very variable and unstable. When the game was reinstated, there was a decrease in VD and this decrease continued across the phase, until Ellie was absent for one day, which was followed by an increase in VD (*M* = 11.51%, range = 0-21.05%). Across the following two phases (B3min and B4min), Ellie's VD remained low and stable. During both phases, she engaged in VD during an average of 5% of observations (range = 0-10%). At the beginning of the B5min phase, VD was low, but there was a subsequent increase during the final data point (*M* = 13.75%, range = 6.67-20.83%).

Katie engaged in a high rate of VD during the initial baseline phase (*M* = 45.83%, range = 30-60%). The researchers noted that much of this talking was due to engagement with another student sitting very close to Katie, but at the next table over. Katie would turn around to talk to her regularly, sometimes standing up in the process. VD decreased immediately and substantially when the CBGG was introduced (*M* = 5.79%, range = 5.56-6.25%). Again, it must be noted that Katie moved seats during this phase due to a request from her classmate with which she agreed. Katie remained in this new seat during the withdrawal phase, and VD immediately increased, albeit not to a level as high as the initial baseline phase (*M* = 21.78%, range = 5-38.89%). There was also a strong decreasing trend across the phase, with the last two data points in the phase being similar to behaviour during the CBGG phase. VD remained low when the game was reintroduced, followed by a large increase for two data points in the middle of the phase (*M* = 13.82%, range = 5-27.5%). When schedule thinning commenced with the B3min phase, Katie's VD remained low and stable (*M*= 3.33%, range = 0-5%). Behaviour remained similar during the B4min phase (*M*= 4.23%, range = 0-7.7%). Katie was only present for one observation session during the B5min phase, and her rate of VD was very high (33.33%).

Out-of-seat Behaviour. Ellie's OOS was variable and downward trending during the initial baseline phase (*M* = 12.5%, range = 0-25%). When the game was introduced, behaviour did not change substantially and trended upwards slightly throughout the phase (*M* = 8.51%, range = 0-16.67%). The withdrawal phase was similarly unstable and there were two data points where no OOS occurred at all towards the end of the phase (*M* = 6.33%, range = 0-16.67%). Out-of-seat behaviour remained low and quite stable during the following game phase (*M* = 3.75%, range = 0-10%). When schedule thinning was introduced across the final three game phases, Ellie's OOS remained low and stable. During the B3min phase, OOS occurred during an average of 1.67% of intervals (range = 0-5%). During the B4min phase, OOS occurred during an average of 5% of intervals (range = 5%). Finally, during the B5min phase, OOS remained at 0% across the two sessions (*M* = 0%, range = 0%).

Katie's OOS was high during the initial baseline phase (*M* = 22.08%, range = 5-33.33%). As previously noted, she would often get out of her seat to talk to a classmate at another table. When the game was introduced, Katie's OOS immediately reduced to 0% and remained at 0% across the whole phase. During the withdrawal phase, Katie's OOS remained low initially, before increasing in the middle of the phase and decreasing again (*M* = 7.11%, range = 0-25%). When the game was reintroduced, OOS remained low and stable (*M* = 1%, range = 0-5%). This trend continued throughout the B3min (*M* = 5.37%, range = 0-11.11%), B4min (*M* = 0%, range = 0%) and B5min phases (*M* = 0%, range= 0%).

Motor Disruption. Ellie's rate of MD was stable at baseline (*M* = 12.08%, range = 10-15%). Introduction of the CBGG saw an initial immediate decrease in MD, followed by an increase, with stabilisation towards the end of the phase (*M* = 5.77%, range = 0-17.5%). Overall, there was a decrease in level when the CBGG was introduced and there was one data point in the phase which overlapped with baseline. When the game was withdrawn, there was a small immediate increase in MD which continued to increase during the first half of the phase. There was then a decrease towards the end of the phase (*M* = 11.22%, range = 5-20%). Motor disruption remained low and stable across the second CBGG phase (*M* = 3.82%, range = 0-5.26%). During the B3min phase, MD was low initially but then increased to a level approaching baseline levels (*M* = 7.04%, range = 0-11.11%). There was a very slight decrease during the B4min phase (*M* = 5%, range = 5%), followed by two contrasting data points in the B5min phase; one very low and one very high (see Figure 5.5.; *M* = 12.5%, range = 0-25%).

Katie's rate of MD was variable at baseline (*M* = 18.33%, range = 5-33.33%), with mostly low rates but with one data point where a very high rate of MD occurred. When the game was introduced, MD decreased immediately and stabilised (*M* = 1.85%, range = 0-5.56%). When the game was withdrawn, MD did not immediately return to baseline levels. It remained low initially, before increasing in the middle of the phase and decreasing again towards the end of the phase (*M* = 12.11%, range = 5-20%). When the game was reintroduced, it did not seem to impact Katie's MD which remained quite high across the phase (*M* = 15.87%, range = 5-26.32%). When schedule thinning commenced, Katie's MD immediately decreased during the B3min phase and remained quite low across the three data points in the phase (*M* = 1.85%, range = 0-5.56%). In the B4min phase, the initial data point reflected a moderate amount of MD, however behaviour decreased during the following two data points (*M* = 3.42%, range = 0-10.26%). In the final B5min phase, Katie was only present for one observation session, and engaged in MD during 6.67% of intervals which reflects a low rate.

Disruptive Behaviour Composite Variable. Ellie's overall rate of DB was high at baseline (*M* = 35%, range = 25-40%) and very similar to whole class DB. There was an immediate decrease initially when the CBGG was introduced, however data were variable and the second data point overlapped with the baseline phase, before DB decreased again (*M* = 20.52%, range = 11.11-30%). Ellie's DB was higher than that of the whole class during this phase.

Despite an initial decrease in DB when the game was withdrawn, Ellie's DB increased sharply towards the middle of phase, before dropping again (*M* = 29.67%, range = 15-55%). Data were variable throughout this phase, however mean DB was similar to that of the class. When the CBGG was reinstated, data stabilised slightly and returned to levels more similar to the previous intervention phase (*M* = 16.51%, range = 10-21.05%). DB remained low and stable across the B3min (*M* = 10.37%, range = 10-11.11%) and B4min (*M* = 10%, range = 10%) phases with little variation in the data. This pattern was sustained in the first data points in the B5min phase, however during the final data point, DB increased to a very high level (*M* = 20%, range = 6.67-33.33%). This coincided with a low level of treatment integrity during the final data point.

Katie's overall rate of DB was very high during the initial baseline phase and always occurred at a higher rate than the class as a whole (*M* = 52.92%, range = 40-66.67%). When the CBGG was put in place, Katie's DB decreased immediately and substantially (*M* = 5.79%, range = 5.56-6.25%) and was lower than that of the whole class. There was no overlap with the initial baseline phase.

Katie's DB increased again when the game was withdrawn with an increasing trend across much of the withdrawal phase (*M* = 32.78%, range = 15-45%). No data points in this phase overlapped with the preceding CBGG phase. The pattern of DB in the following CBGG phase was not as stable as in the previous CBGG phase and Katie's DB remained higher than the whole class on average. DB was variable and there was a lot of overlap with the preceding withdrawal phase (*M* = 23.53%, range = 5-42.12%). During the B3min phase, DB was initially quite low with an upward trend across the three data points (*M* = 10.56%, range = 5-16.67%). In the following B4min phase, DB remained low and followed a decreasing trend across the phase (*M* = 6.79%, range = 0-15.38%). Katie was present for one of two observation sessions in the B5min phase and her DB was very high (40%; data point 26).

Figure 5.3.

Percentage of Intervals with AEB and DB Across Study Phases for Ellie Superimposed over Whole Class Data for Comparative Purposes

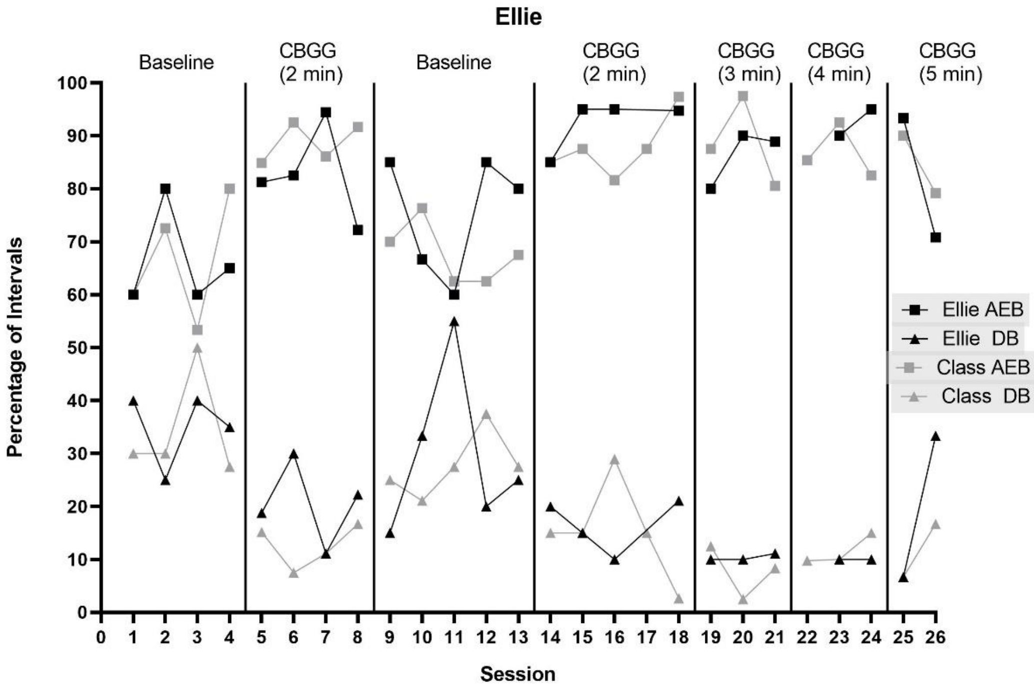


Figure 5.4.

Percentage of Intervals with AEB and DB Across Study Phases for Katie Superimposed over Whole Class Data for Comparative Purposes

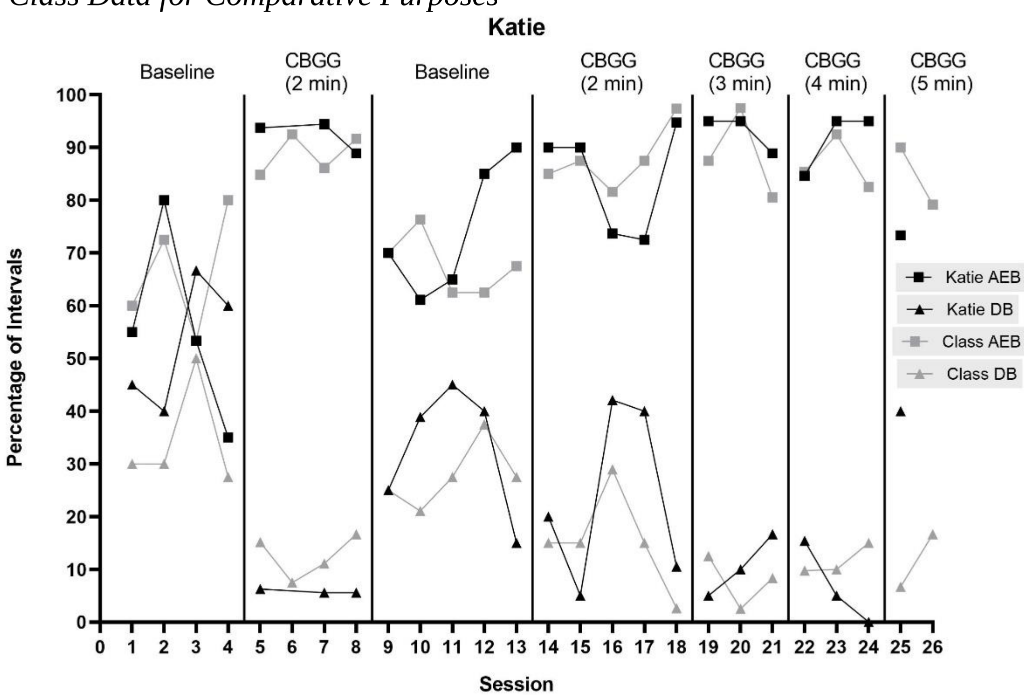


Figure 5.5.

Percentage of Intervals with VD, OOS and MD Across Study Phases for Ellie

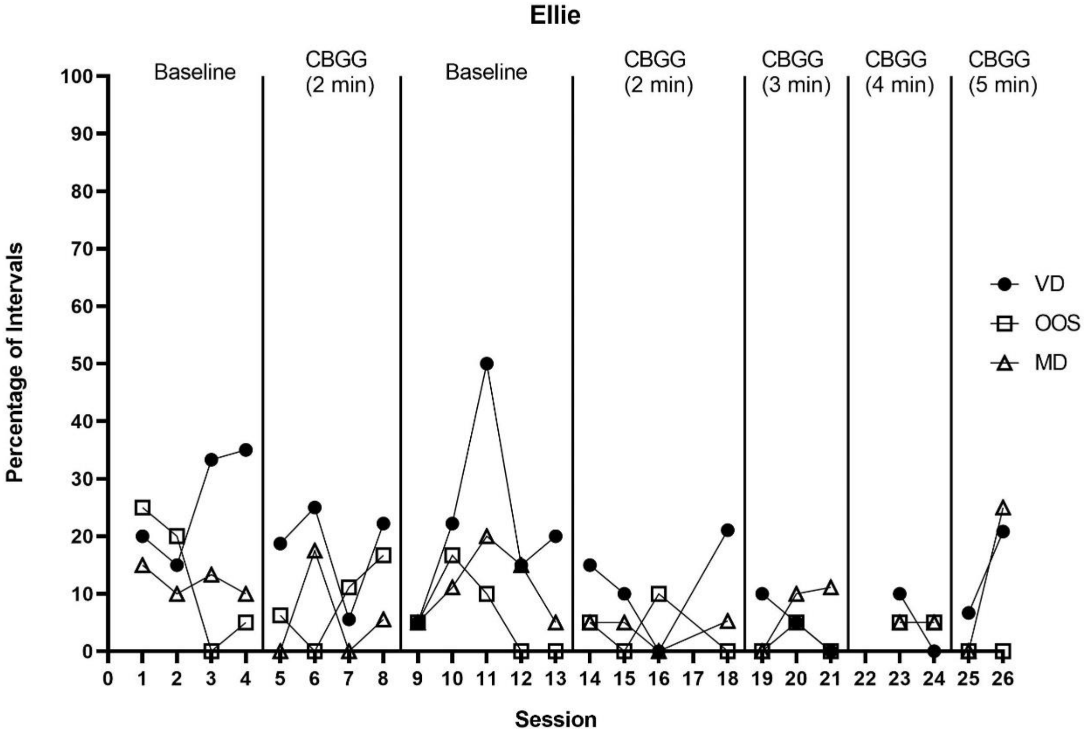
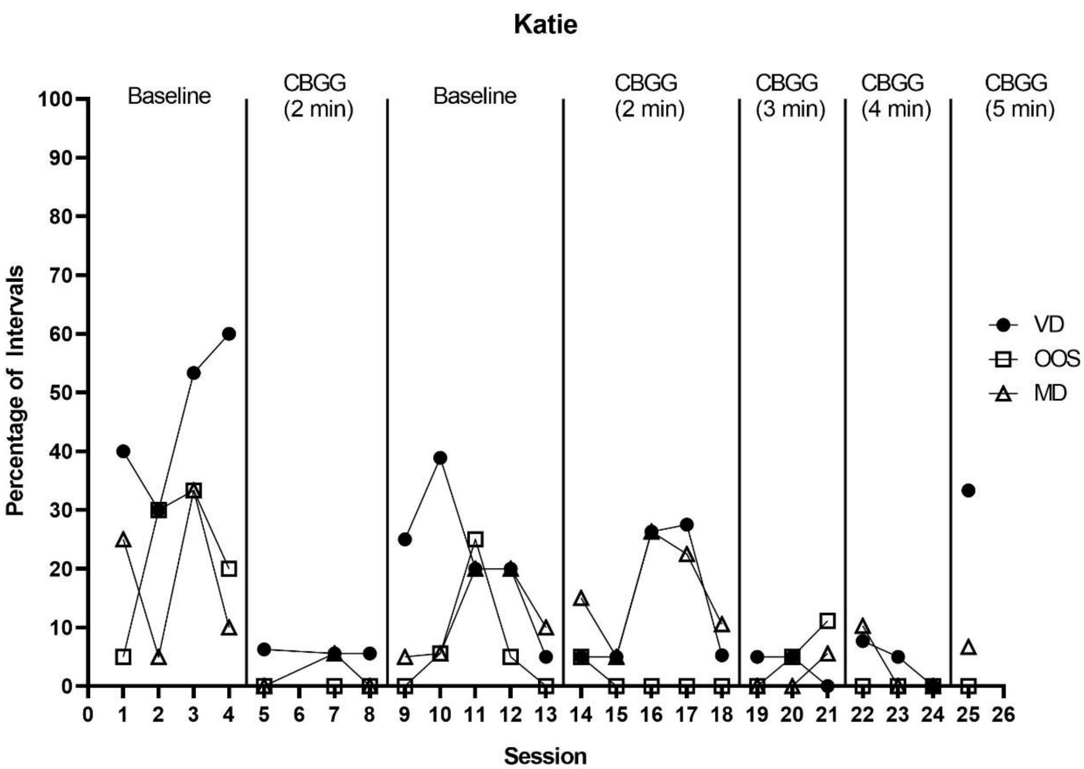


Figure 5.6.

Percentage of Intervals with VD, OOS and MD Across Study Phases for Katie



Effect Sizes for Behaviour Data

Tau effect sizes were calculated for each phase change and weighted mean effect sizes were subsequently calculated for each participant and outcome. Table 5.4. provides an overview of these effect sizes. To calculate an effect size for the schedule thinning phases, each of the three phases were compared separately to the second baseline phase (i.e., the withdrawal phase/A2). In calculating the weighted means, the phases in which thinning was introduced were not considered, therefore the weighted means provide an effect size for the initial A-B2min-A-B2min phases only.

Table 5.4.

Tau Effect Sizes Across Phase Changes for the Whole Class and Target Students

Outcome	Tau A1-B2min(1)	Tau A2-B2min(2)	Tau A2-B3min	Tau A2-B4min	Tau A2-B5min	Weighted Mean Tau[a]
Whole Class						
Verbal Disruption	-.77	-.518	-.76	-.745	-.707	-.67
Out-of-seat Behaviour	-.481	-.274	-.105	-.385	-.149	-.38
Motor Disruption	-.481	-.732	-.732	-.596	-.69	-.65
Total Disruptive Behaviour	-.77	-.531	-.745	-.745	-.707	-.67
Academically Engaged Behaviour	.756	.762	.745	.745	.707	.76
Ellie						
Verbal Disruption	-.283	-.34	-.596	-.552	-.138	-.31
Out-of-seat Behaviour	-.144	-.169	-.338	0	-.49	-.16
Motor Disruption	-.392	-.49	-.304	-.490	0	-.45
Total Disruptive Behaviour	-.674	-.46	-.745	-.707	-.212	-.58
Academically Engaged Behaviour	.674	.712	.506	.707	.283	.7
Katie						
Verbal Disruption	-.775	-.156	-.685	-.608	.359	-.55
Out-of-seat Behaviour	-.816	-.446	0	-.548	-.387	-.68
Motor Disruption	-.645	-.244	-.62	-.456	-.12	-.23
Total Disruptive Behaviour	-.775	-.241	-.634	-.634	.239	-.57
Academically Engaged Behaviour	.756	.463	.646	.547	.115	.63

[a]This weighted mean does not consider the phases where the CBGG schedule was thinned (i.e., B-3min, B-4min and B-5min)

Social Validity

Teacher Social Validity

Teacher feedback collected via self-report on the BIRS (Elliott & Von Brock Treuting, 1991) was positive. She scored the game 90 out of a possible 90 on the acceptability subscale, responding with ‘strongly agree’ to all statements. On the effectiveness subscale, she scored the game 35 out of a possible 42. Her mean rating on this scale was 5 (range = 4-6). The teacher’s rating on the efficiency subscale was 11 out of a possible total of 12. Her mean rating on this scale was 5.5 (range = 5-6).

Ms. Ellis was asked to provide additional comments on the schedule thinning element of the game and some general comments. When asked whether she had a preference for any particular version (2, 3, 4 or 5 min between points), she responded as follows: *“The 2/3 minute intervals was intense and meant individual help was interrupted”*. When asked about her perceptions of effectiveness across the different time intervals, she responded as follows: *“Shorter intervals kept children's attention on whether they were getting rewards in the beginning but for the long term use [of] 5 min intervals is more manageable”*. She also stated that the 5 min version of the game was *“easier to implement”*. When asked for additional general comments, Ms. Ellis stated that *“the CBGG worked very well in [her] classroom”*. She also stated that sometimes she *“found it difficult to be consistent with the timetable but this only applies to the research. A teacher using this game can be more flexible”*. Finally, she expressed some concern over her objectivity in awarding points.

Student Social Validity

Nineteen students completed the modified CIRP (Mitchell et al., 2015; Witt & Elliott, 1985). The mean rating across the survey was 7.32 (range = 5-8), indicating a positive perception of the game overall. When students responded negatively to an item (i.e., answered ‘no’ to items 1-6 or ‘yes’ to items 7-8), the student researcher asked them why they had answered this way and wrote down their response. For example, one student stated that there were things his friends did not like about the game, and when asked why, responded stating *“when they lose and don't get stamps”*. Another student stated that there were things they did not like about the game and when asked why, responded stating “*When friends distracted me”*. One student who thought the game was not fair expanded, stating that he was *“out of his seat only for a minute”* and that could have cost his team a point.

Discussion

The current study aimed to evaluate the effectiveness of the CBGG with a senior infant population and if the effectiveness of the game could be sustained over several phases where the schedule of reinforcement was thinned. Individual and group behaviour was monitored to assert whether particularly disruptive students responded differently to the game compared to the whole class group. The results of the current study support the efficacy of the CBGG in increasing AEB and decreasing three types of DB in a mainstream, senior infants class. These findings are in line with research on other, similar populations (e.g., Lynne et al., 2017; Tanol et al., 2010). The CBGG with 2 min intervals between reinforcement opportunities produced significant increases in AEB and decreases in DB across the class across two intervention phases, and the subsequent schedule thinning procedure demonstrated that the game could remain effective with longer intervals between reinforcement opportunities. The CBGG was also effective in targeting AEB and DB in two individual target students, although effects were not as strong as for the whole class. Results will be discussed in terms of each research question.

Effectiveness of the CBGG with a Senior Infants Class

The CBGG was effective in targeting AEB and DB across the whole class group with a dense schedule of reinforcement (2 min). Immediate and large increases in AEB were evident each time the CBGG was introduced. The CBGG produced large weighted mean effect sizes on AEB (.76) and DB (-.67), with large weighted mean effect sizes also detected for two sub-categories of DB: VD and MD. These findings align with previous research which demonstrated the efficacy of the CBGG with young primary school students. For example, Tanol et al. (2010) investigated the CBGG with individual students across two kindergarten classes, finding that it led to reductions in rule violations. In that study, the teachers were responsible for determining the rate of reinforcement and were not prompted to carry out ‘behaviour checks’ at fixed intervals. This made the schedule of reinforcement variable across the 10 min game session. The present study manipulated the rate of reinforcement delivered by the teacher by establishing schedules. Lynne et al. (2017) also demonstrated that the CBGG could be an effective intervention with a young primary school aged population (i.e., a first grade class), but again, the rate of reinforcement appeared to be determined by the teacher, in that they decided themselves when and whether to award points, rather than being prompted to do so. The current study saw the CBGG applied with more structure, similarly to how it was applied in studies by Wright & McCurdy (2012) and Wahl et al., (2016). Wright and McCurdy (2012) demonstrated the efficacy of the CBGG with a kindergarten class on a VI4min schedule and Wahl et al., (2016) produced similar results in four young primary school classes on a VI5min schedule. Although similar, the present study focused on establishing control over behaviour with a dense fixed schedule in the initial phases, before thinning the procedure gradually. Wright and McCurdy only implemented the CBGG in one study phase (ABAC design) and Wahl et al., did not incorporate a withdrawal phase. The present study addressed these shortcomings with three attempts to demonstrate an intervention effect (WWC, 2017; 2020) before introducing a thinning procedure.

Effectiveness of the CBGG with Individual Target Students in Senior Infants

The teacher's two chosen target students responded quite positively to the CBGG, however not to the same extent as the wider class group. In general, the target students' levels of AEB and DB increased and decreased respectively when the CBGG was in place compared to baseline/withdrawal phases. The proportion of overlap between phases was greater than that of the whole class. The CBGG produced large weighted mean effect sizes for AEB and moderate weighted mean effect sizes for DB across both students. The CBGG produced very similar effects on the target students' DB (-.58 and -.57 respectively), and the game produced a slightly larger effect on AEB for Ellie (.7) compared to Katie (.63). As noted in the results section, Katie's initial improvements in behaviour must be analysed considering the identified confounds related to the seating plan. Although it is evident in the data that Katie's behaviour was influenced by changes in the seating plan, the initial change (at the beginning of the first B phase), was at the request of another student upon hearing about the game. This was a request that Katie subsequently agreed with and Ms. Ellis permitted the seat change. Katie willingly moved her seat to the other side of the table on every subsequent day during that phase when the game was in place. It was clear that the game appeared most effective for Katie when considering the transition from phase A1-B1 (i.e., coinciding with the change in seats), rather than between phase A2-B2. The seat change was out of the control of the researcher, however has implications for the interpretation of the data. This demonstrates one of the benefits of the single-case research design and collecting continuous data on student behaviour, in that the impact an environmental confound has on behaviour can be visualised in the graphed data. It does mean however, that it is difficult to separate the effects of the CBGG and the seat change. The results demonstrate how individual student behaviour can be sensitive to small changes in the environment such as the seating plan, whereas the whole class behaviour can remain quite stable under these changes.

The findings from the two individual students lend support to the idea that although an intervention may appear to be highly effective across a whole class group, individual students may not perform as well in certain areas. For example, Ellie's VD was quite high at baseline and occurred at a mean level identical to that of the whole class (M = 25.83%). When the CBGG was put in place, mean VD across the class decreased by 20.92%, whereas mean VD for Ellie decreased by just 7.95%. Previous studies have identified non-responders to the GBG in a similar way. Donaldson et al. (2017) collected data on 12 individual students as part of an evaluation of the GBG with individuals and found that 3 of the 12 students did not respond to the GBG. They could therefore be potentially classified as 'non-responders' and be referred for more intensive behavioural support. Neither target student in the present study was a total non-responder and both benefited in some way from the CBGG.

Maintained Effectiveness with Schedule Thinning

The reinforcement schedule (i.e., the teacher behaviour check schedule), was initially thinned to 3 min intervals after the A-B2min-A-B2min phases were complete. Across the whole class, effects on AEB were largely maintained across the three sessions, with no overlap between this phase and any previous baseline phases. The same was true for whole class DB. Ellie's AEB and DB remained similar to the preceding B2min phase during the B3min phase. Katie's behaviour was slightly more variable in that there was an upward trend in DB during the B3min phase. Nonetheless, her level of DB remained much lower than during previous baseline/withdrawal phases. This provided enough initial evidence to continue with the B4min phase. This phase produced very similar results across the whole class when compared with the B3min phase. AEB remained high with no overlap, and DB remained low with no overlap. Ellie's AEB remained high and stable and DB remained low and stable. Katie's AEB was also high across this phase and her DB decreased across the phase. In the B5min phase, AEB remained high and DB low initially, but a decrease in AEB and increase in DB were apparent during the final data point. A very similar pattern of behaviour was evident for Ellie. This may be attributable to low teacher treatment integrity during this data point. During this school week, teachers were particularly busy preparing for rumoured impending school closures due to the Covid-19 pandemic, and Ms. Ellis was notably concerned with this. School closures were subsequently announced two days after this data collection session. Katie was only present for one data collection session during the B5min phase, and her behaviour deteriorated when compared to the preceding B4min phase. Strong conclusions cannot be drawn based on just one data point however. Overall, the reinforcement schedule was successfully thinned to at least 4 min and tentative conclusions can be drawn for the effectiveness of the CBGG with 5 min intervals. The CBGG with 5 min intervals was applied during a week in which many competing environmental influences were at play, particularly for the teacher.

The above discussed results from the schedule thinning phases support the idea that schedule thinning is a potential solution to lessening the workload for the teacher during CBGG implementation. Although the CBGG has been applied with variable schedules (Wahl et al., 2016; Wright & McCurdy, 2012), dense schedules (Ford, 2017) and teacher determined schedules (Tanol et al., 2010), no previous study had looked at beginning with a dense schedule and thinning it over time. No previous research has looked at the concept of schedule thinning during the CBGG, however, other research has demonstrated the efficacy of schedule thinning during DRO for other behaviours (e.g., Bergstrom et al., 2011). Future research should perhaps consider examining within session schedule thinning. This would involve the teacher playing the game with a dense schedule initially (e.g., 2 min intervals) and progressing to larger intervals within the same session/day. This may be a potential solution for teachers who are concerned with behaviour reverting quickly to baseline levels when the game is withdrawn (Donaldson et al., 2015). This issue was raised by Ms. Daly, the teacher in the study outlined in Chapter 4. Schedule thinning within the same game session as a method of maintaining behaviour over a longer game period is therefore a potential area for future research.

Teacher Social Validity Rating

Ms. Ellis found the CBGG to be highly acceptable, effective, and efficient, with mean scores of 6, 5 and 5.5 on these sub-scales of the BIRS (Elliott & Von Brock Treuting, 1991) respectively. The ratings she gave were higher than those given by teachers in a recent study with a similar population of students where the CBGG was applied with Class Dojo technology (Lynne et al., 2017), and higher than the ratings given by Ms. Daly in Chapter 4. The teacher in the study by Lynne et al., (2017) rated the game with a mean of 5.24 for acceptability, 4.71 for effectiveness and 5 for time of effectiveness/efficiency. There were two key differences in the methods of implementing the game described by Lynne et al., and those described here. The CBGG was implemented with the incorporation of Class Dojo by Lynne et al., whereas in the present study, a more traditional scoreboard was used where the teacher simply coloured in points when awarding them. As previously mentioned, Lynne et al., did not report that a schedule was set for the teacher to conduct behaviour checks and award points. The teacher in the present study noted that the game was easier to implement as the schedule was thinned but appreciated that the denser schedule kept the children’s attention. It is evident that she could appreciate the benefits of beginning with a dense schedule and thinning it over time, rather than criticising the dense schedule. When comparing the ratings with Ms. Daly’s ratings in Chapter 4, Ms. Ellis rated the game slightly higher across all three scales. Notably, she rated the game highly on the effectiveness scale (M = 5) whereas Ms. Daly rated the game as slightly ineffective (M = 3.29). As previously noted, the questions around effectiveness of the game on the BIRS mainly referred to potential for generalisation and maintenance of the game’s effectiveness rather than its effectiveness during intervention sessions [e.g., “The intervention will produce a lasting improvement in the students’ behaviour”; “The students’ behaviour will remain at an improved level even after the intervention is discontinued”; “Using the intervention should not only improve the students’ behaviour in the classroom, but also in other settings (e.g., other classrooms, home)”]. It is possible that due to the schedule thinning procedure, Ms. Ellis could appreciate the potential for the maintenance of the game. It is also important to note that although the CBGG was applied in a similar fashion across both studies, Ms. Ellis’ class were much younger than Ms. Daly’s (senior infants versus fourth class) and were mixed-sex rather than single-sex boys. Ms. Ellis also noted that there were some difficulties in being consistent with the timetable. During the research process, she tried to consistently do a mathematics class at the same time each day, which was evidently difficult at times. In a primary school classroom, teachers spend the entire day with the same class group and therefore have the benefit of being able to choose when to do each activity. It was requested that the teacher do a similar activity during observation sessions daily and therefore she attempted to stay consistent with the mathematics class.

Student Social Validity Rating

The 19 students who completed the social validity measure responded very positively to the CBGG, rating it a mean of 7.32 across 8 items. This is the highest rating given to the CBGG across all studies in this thesis, however the students in Chapter 4 did not have the opportunity to rate the game. Although some students in Ms. Ellis’ class reported not liking elements of the game, when probed further, these were usually related to incidences where they didn’t receive a point or when they didn’t receive enough points to get a prize, rather than the mechanics of the game itself. Some studies on the CBGG in recent times have not obtained student social validity measures (e.g., Lynne et al., 2017;

Tanol et al., 2010). This includes the study in Chapter 4 of this thesis, which demonstrated how unexpected and uncontrollable events can impact elements of the planned research process. Students in the CBGG/GBG comparison study by Wahl et al., (2016) rated the CBGG positively and most students preferred the CBGG to the GBG. The kindergarten students in the study by Wright & McCurdy (2012) also rated the CBGG positively and rated it slightly more positively than the GBG. The results in the present study cannot be directly compared to the aforementioned studies, as slightly different ratings scales were used, however, overall, it appears that the CBGG is a well-liked intervention among young primary school students.

Limitations

The results of the current study must be considered in light of a number of limitations. Firstly, as with other studies discussed in this thesis, only one classroom in one school setting was recruited. Although it was planned that other classes would be recruited, this was not possible during this project. This limits the generalisability of the results. A second limitation relates to some confounding variables which could not be avoided during the study. The most notable of these was the change in seating plan on occasion throughout the study (individual student changes and whole class changes). Importantly however, seating plan changes are a regular occurrence in school classrooms, with teachers often changing their seating plan monthly. It would have been unfair to request the teacher to maintain the same seating plan through the course of this project as it spanned three months and only involved 20 min of the longer school day. Seating changes appeared to have the greatest impact on Katie's behaviour and in places it is difficult to separate the effects of a seating change and the CBGG. Another limitation was that treatment integrity, although relatively high on average, was low on occasion throughout the study, particularly during the very last data point. This limits conclusions which can be drawn for the CBGG with 5 min intervals between reinforcement opportunities, as it may have been more effective if treatment integrity had been higher. Unfortunately, data collection was ceased unexpectedly and earlier than planned, so additional data points could not be collected in the B5min phase. Another limitation relates to the method of choosing the target students. The students were chosen based on teacher evaluation of the class. The teacher was asked to consider the rate of disruption and student attendance when choosing target students (i.e., to choose those with a good attendance record so that enough data could be collected on their behaviour to draw reliable conclusions). A more systematic method, such as collecting baseline data on all students, may ensure that the most objectively disruptive students are chosen for individual monitoring. Both Ellie and Katie displayed rates of AEB lower than and rates of DB higher than the rest of the class at baseline however, indicating that they were suitable candidates for individual behaviour monitoring. Finally, a schedule thinning procedure was used here and started with 2 min intervals between opportunities for points and progressed to 5 min intervals. It is possible that the class may have responded well to a thin schedule of reinforcement in the first instance. Future research could perhaps be more objective in choosing a starting point for intervals by probing the game with differing interval lengths or calculating the mean interresponse time between incidences of DB of the most disruptive students. Furthermore, the schedules used here were similar to those used in other studies and therefore a greater contribution to knowledge may be made by thinning schedules even further. There was a risk that starting with a very thin schedule would lead to null intervention effects, which may have led to low social validity perceptions from the teacher. In order to increase the potential for intervention effects and thereby teacher buy-in, the choice was made to begin with a dense schedule, and thin in 1 min increments. Nevertheless, the current research provides evidence for continued efficacy of the CBGG with successively thinner reinforcement schedules over time.

Implications for Future Research and Practice

The current study provides further evidence for the use of the CBGG in primary school classrooms, particularly those at the lower end, building on results from Chapter 4. It is also one of few studies demonstrating the effectiveness of the CBGG in an Irish primary school setting (with Chapter 4 being the only other study with Irish primary school children). Although the evidence is preliminary, taken with other evidence, teachers may consider this game for use in their classrooms. The study has demonstrated the game's efficacy across a whole class group but has also demonstrated that not all individual students respond in the same way to the CBGG. This suggests that teachers should be mindful of a) students who do not respond as well to class wide interventions and b) total non-responders. Ellie and Katie would not be classed as total non-responders in this study and the CBGG did have some positive effects for each of them respectively, however, future research may investigate ways to intervene when individuals do not respond as well as the whole group. The evidence here for schedule thinning as part of the CBGG is promising and provides a starting point for further research on the concept. Future research may evaluate whether the CBGG is effective initially with a dense schedule, and then prolong the game with a thinned schedule during the same session. This may provide a solution for teachers who are concerned that disruptive behaviour becomes an issue immediately after game cessation, such as the teacher in Chapter 4. It may also be a worthwhile avenue to examine if the CBGG can be effective with even thinner schedules of reinforcement beyond the 5 min examined in the current study. For example, Bergstrom et al., (2011) successfully thinned a schedule from 10 s to 600 s (10 min) in a previously discussed study. This was a 5900% increase in seconds. The current study implemented a 150% increase from 2 min (120 s) to 5 min (300 s). There is therefore potential to further extend the examination of schedule thinning in the context of the CBGG.

Conclusion

Overall, the current study has provided further good quality evidence for the CBGG with a lower primary population. The concept of schedule thinning as part of the CBGG has been introduced, however requires further examination in different contexts. Individual students responded well but differently to the whole class during this iteration of the CBGG. Therefore, the benefits of collecting data on individual students as well as the whole class group cannot be ignored.

Chapter 6: Discussion

This final chapter of the current thesis will summarise and synthesise the findings of the programme of research, maintaining focus on the specific research aims and objectives. Study results have already been discussed at the end of each chapter, however this chapter will integrate the findings, consider them more broadly within the literature, and speak to the broader implications for future research and practice. Strengths and limitations of the research programme will be outlined, followed by concluding statements.

Summary of the Programme of Research

The primary aims of the current programme of research were to understand and synthesise the body of literature around behavioural game-based classroom management interventions in mainstream classrooms and to evaluate some of the features of game-based classroom management across diverse Irish classroom settings. With these aims in mind, research objectives involved the development and completion of a systematic literature review on game-based classroom management interventions in mainstream classrooms, the identification of an intervention to investigate within Irish classroom contexts and the evaluation of this intervention across a diverse range of contexts. The need to evaluate these types of interventions in Irish settings was clear, given the diversity present in the school system and discrepancies in types of discipline systems used across schools (e.g., Smyth & Quail, 2017), different rates of engagement in boys versus girls (e.g., McCoy et al., 2012) and in DEIS versus non-DEIS schools (e.g., Darmody et al., 2008). Given the range of previous research on the GBG and CBGG in diverse settings, it was expected that the CBGG would be effective in Irish classrooms, however evaluating this experimentally was an important avenue for research, not least because teachers may be more likely to adopt a strategy if they see that it has been implemented successfully with a classroom similar to their own (Joram et al., 2020).

Chapter 2 was a review designed to synthesise existing research in the broad area of game-based classroom management. Studies identified for inclusion were subject to strict design standard evaluations (WWC, 2017; 2020) and data on study characteristics were extracted from those studies meeting the design standards. The review provided an overview of the settings in which and populations with which game-based interventions had been applied, developed a list of the game-based interventions which had been applied and evaluated the effectiveness of these games by calculating effect sizes and conducting a series of meta-analyses.

Chapter 3 consisted of two studies designed to evaluate the effects of the CBGG in three secondary school classrooms. In both studies, visual feedback was manipulated such that students were exposed to a version of the game both with (CBGG-i) and without (CBGG-d) immediate visual feedback. Study 1 employed a reversal/withdrawal design (ABACABAC) across one class group. Study 2 expanded upon this by allowing for counterbalancing of intervention conditions during a reversal/withdrawal design across two class groups (ABACABAC versus ACABACABAB). In both studies, student AEB and DB were monitored across study phases and in Study 2, teacher statements were monitored as well as student behaviour to assert whether the CBGG influenced rates of teacher praise and/or reprimands.

Chapter 4 investigated the CBGG in a single-sex (all-boys) school setting, a type of setting neglected in previous CBGG research. A reversal/withdrawal design with phases ABAB was applied. AEB and DB of two individual target students as well as of the whole class group were monitored to assert whether differential intervention effects would be present for individuals compared to whole class data.

Chapter 5 aimed to establish whether the schedule of reinforcement could be thinned successfully over time during CBGG implementation, while maintaining its effects on AEB and DB. In other evaluations of the CBGG, the schedule of reinforcement is generally static or set to a variable schedule (e.g., Wahl et al., 2016; Wright & McCurdy, 2012). The study was conducted with a senior infant, mixed-sex class. Like Chapter 4, behaviour of two individual target students and the whole class group were monitored during a reversal/withdrawal design with phases ABAB-B3min-B4min-B5min. The schedule of reinforcement was thinned from 2, to 3, to 4, and eventually to 5 min.

Synthesis of Key Findings

When discussing the key findings, it is important to look collectively at the range of studies conducted and evaluated across this thesis. Initially, 395 individual experiments were evaluated across 49 studies included in the initial stages of the systematic review in Chapter 2. Detailed coding took place for 169 of these experiments which met the WWC design standards (WWC, 2017; 2020) across 30 studies and graphed data were extracted from 143 experiments in order to facilitate effect size calculations. Such a detailed review of game-based classroom management literature had not been conducted prior to this and Chapter 2 situates itself in a unique literature gap between general evaluation of behavioural strategies for classroom management (e.g., Chaffee et al., 2017) and reviews of specific game-based interventions for classroom management, such as the GBG

(Bowman-Perrott et al., 2016). Across studies and outcomes (positive and disruptive classroom behaviours) included in the review in Chapter 2 a majority of the effect sizes calculated were large, providing a strong platform for further investigation of these game-based interventions, with a focus on procedural variations and diverse populations.

In a follow-up to the Chapter 2 review, and some of the gaps in the literature noted in it, the CBGG was trialled across five Irish mainstream classrooms (Chapters 3-5). The game was found to be effective in targeting AEB and DB from first-year secondary school students (adolescents) down to senior infant primary school students. If an experiment is characterised as it was during the review in Chapter 2 (i.e., one participant case and one outcome in a reversal/withdrawal design), then 40 individual experiments formed part of the empirical chapters (3-5) in this thesis. Thirty-six of these were experiments examining the effects of the CBGG on some student target behaviour, with the remaining 4 evaluating the impact of the CBGG on teacher praise and reprimands. The 36 empirical experiments on student behaviour build significantly upon the secondary evidence evaluated in Chapter 2. Forty-two weighted mean effect sizes were calculated across these experiments (note that two weighted means were calculated for each outcome during Chapter 3 as two different versions of the game were examined). Of these 42 effect sizes, 25 were large to very large, 15 were moderate and 2 were small. The two small effect sizes refer to the effect of the CBGG on the OOS of two target students; Ben in Chapter 4 and Ellie in Chapter 5. Evidently, OOS was not a huge issue for Ben, whereas Ellie's OOS remained very variable across study phases. The findings across studies will be discussed in more detail in the next section, but to summarise briefly here, it is apparent that across the studies conducted as part of this thesis, the CBGG was an effective intervention which had a positive impact on AEB and led to reductions in student DB. These findings align with other previous research which demonstrated the efficacy of the CBGG with mainstream students across primary and secondary school levels (Ford, 2017; Groves & Austin, 2020; Lynne et al., 2017; Wahl et al., 2016; Wright & McCurdy, 2012). The findings also expand upon previous research by examining procedural variations of the game (e.g., visual feedback, schedule thinning, weekly prizes only) and examining the game with diverse populations (e.g., in an Irish context, with 'at-risk' adolescents, in a single-sex boys setting). Overall, the range of experiments conducted support the efficacy of the CBGG with procedural variations in Irish classrooms.

Based on evidence from the review in Chapter 2, it can be asserted that these studies are the first of their kind to be conducted in Ireland. Prior to this, no empirical investigation of game-based interventions for classroom management had taken place with Irish students. This fills an important gap in the literature, given most game-based interventions have only been evaluated in a US context. By targeting ‘at-risk’ students (by school DEIS status) in Chapters 3 and 4, the results provide support for use of the CBGG with some of the populations most in need of support in an Irish context. The promising results have identified evidence gaps and provided a premise for a whole range of future research. We will now turn to consider the key findings of this thesis individually, namely: **1)** the effectiveness of the CBGG on student behaviour across both individuals and groups, **2)** effectiveness of the CBGG maintained with significant procedural variations including feedback manipulation, weekly prizes only and schedule thinning, **3)** the impact of the CBGG on teacher behaviour and **4)** the social validity of the CBGG.

The Effects of the CBGG in Irish Schools

The findings of the empirical studies outlined in the current thesis have supported the efficacy of the CBGG in targeting AEB and DB across diverse student populations. The systematic review outlined in Chapter 2 provided an initial insight into the potential for this intervention, with evaluations by Fallon et al., (2018), Ford (2017), Hernan et al. (2019), Lynne et al., (2017), Wright and McCurdy (2012) and Tanol et al., (2010) meeting WWC design standards (with or without reservations; WWC, 2017; 2020) and reporting positive effects of the game from populations ranging from kindergarten to ninth grade. Patrick et al. (1998) also evaluated a version of the CBGG where response-cost procedures were incorporated (i.e., points could be rescinded as well as awarded). Behaviours targeted across these studies included various types of disruptive behaviour, academic engagement, mobile phone presence during class and appropriate/inappropriate social behaviour during physical education. Inconsistent terms had been used to label the CBGG, however all of the studies on the positive GBG for which effect sizes could be calculated in the review, produced moderate to large effects on behaviour. Given the recency of these studies (i.e., most have been published in the last five years), the review has assisted in identifying the surge in research on the CBGG. This surge is not surprising given the move towards positive behaviour strategies in schools. The GBG as originally applied with its focus on occurrences of disruption may not align with modern recommendations in school behaviour policies. For example, in an Irish context, school behaviour policies should encourage positive behaviour as much as possible (National Educational Welfare Board, 2008). This surge in research on the CBGG, paired with the focus on positive behaviour strategies in Irish schools, made it an ideal candidate to evaluate during the empirical chapters in this thesis.

Positive game-based classroom management interventions had not yet been evaluated in Irish populations (see review in Chapter 2) and therefore the CBGG was chosen for further evaluation based on its positive method which would align with current recommendations and practices in Irish classroom management. This positive and proactive focus which is characteristic of the CBGG, gives teachers a potential tool with which to pre-plan for the occurrence of disruptive behaviour and counter-act it, rather than waiting for incidences of disruptive behaviour to occur and reacting to it. In general, a move from punitive to more positive approaches to classroom management has been suggested in the literature (e.g., Sugai & Horner, 2006), and this has been supported by research findings that use positive behavioural strategies are positively associated with dimensions of school climate such as student-teacher relationships (Mitchell & Bradshaw, 2013). The CBGG demonstrated positive effects on student behaviour across the four studies and five classrooms in this thesis, with each study evaluating different aspects of the game or focusing on a different student population. Visual feedback, individual versus class wide efficacy and schedule thinning were unique focuses across the four studies. The populations under study ranged from an early primary school mixed-sex class (senior infants), to a single-sex fourth class group, to at-risk adolescent students (first years). Thirteen of the sixteen effect sizes calculated for the effects of the CBGG on class-wide behaviour (AEB and DB) were large. The other three effect sizes were moderate. Across the four individual target students in two of the classrooms (Chapter 4 and 5), effect sizes were also all moderate to large for AEB and DB. These findings align with large effect sizes uncovered in Chapter 2 for the impact of the CBGG on adolescent (Ford, 2017) and primary-aged children's' behaviour (Lynne et al., 2017; Tanol et al., 2010) in school settings in the USA.

Interestingly, the moderate to large weighted mean Tau values identified across Chapter 3-5 also align closely with recent studies conducted on the GBG. It is important to make this comparison as the CBGG, while similar to the GBG, maintains a focus on positive behaviour. There is a large body of evidence in support of the traditional GBG, as evidenced in Chapter 2 and other reviews (e.g., Bowman-Perrott et al., 2016; Chaffee et al., 2017), while less research has evaluated the CBGG. Large effect sizes were identified in Chapter 2 for the impact of the GBG on adolescent behaviour (Kleinman & Saigh, 2011; Mitchell et al., 2015) and primary-aged students' behaviour (Dadakhodjaeva et al., 2019;

Donaldson et al., 2015; Nolan et al., 2014). The comparability of the results from Chapters 3-5 to recent research on the traditional GBG has important applied implications. As outlined earlier, in an Irish context, it is seen as important to choose a positive behaviour strategy where possible to promote good behaviour. The National Educational Welfare Board and Tusla (the Child and Family Agency in Ireland) prepared a document outlining guidelines for schools in developing a code of behaviour and this document emphasises that behaviour codes should give priority to promoting and affirming good behaviour (National Educational Welfare Board, 2008). The National Council for Special Education's behaviour support service also lists promotion and support of positive behaviour as one of their guiding principles (National Council for Special Education, 2020). For classroom management interventions to fit with modern recommendations, school ethos and policy, it is imperative that they also maintain a positive focus. This research therefore fits well with recommendations by state departments and the current behavioural focus in Irish schools and provides appropriate evidence for a strategy which teachers may consider for more widespread use.

Recent research examining factors that influence teacher's likelihood to implement research-based practices in the classroom identified the teacher's assumptions about the similarities between a study population and their own students as an important factor in their decision (Joram et al., 2020). This reiterates the importance of evaluating strategies like the CBGG with Irish populations, even when there are many studies supporting its efficacy with students in the USA. The collection of data on individuals and groups of students in Irish school contexts situates this research in a gap which was previously neglected. The studies presented in the current thesis represent a hugely important development in Irish classroom management research, given the individuality of each classroom in the country. Even though a small number of classrooms and students were recruited here, each was unique. The secondary school classes that were recruited in Chapter 3 were mixed-sex, at-risk groups (according to the school DEIS status). Similar second-level classes have been exposed to the GBG and CBGG in US contexts (e.g., Hernan et al., 2019; Kleinman & Saigh, 2011), but never in Irish contexts. Chapter 4 involved a single-sex boys' school setting at primary school level, which although commonplace in Ireland, may not be the norm in other countries. This population were also deemed at-risk based on the school's DEIS status. To enhance the likelihood of Irish teachers adopting this approach, it was important to examine the effectiveness of the CBGG in these single-sex contexts, particularly single-sex boys' settings where punitive strategies may be common (Smyth & Quail, 2017). The effectiveness of the GBG and CBGG has been well-established in lower primary school settings (e.g., Dadakhodjaeva et al., 2019; Donaldson et al., 2018; Lynne et al., 2017; Tanol et al., 2010; Wright & McCurdy, 2012), such as the class included in Chapter 5, but again, it is important to note neither the GBG nor CBGG had been applied with such a group in an Irish context previously. Although the senior infant population recruited here were not deemed ‘at-risk’ (i.e., the school did not maintain DEIS status), the teacher approached the student researcher upon hearing about the study and high rates of disruption were evident in the class. Taking all of these considerations into account, Irish teachers may be more inclined to adopt the CBGG as a classroom management intervention given its demonstrated efficacy with Irish students.

Effects on Individual Students. The review in Chapter 2 presented evidence for game-based interventions on group and individual behaviour. The GBG, CBGG, CW-FIT, the Praise game, a self-management game and the KBCO game were all evaluated with individual students. In some studies, group and individual monitoring of behaviour took place (e.g., Dadakhodjaeva et al., 2019) and in others, only individual behaviour was monitored (e.g., Hine et al., 2015). Chapters 4 and 5 differentiated between individual and group behaviour by collecting data on two teacher-selected target students in each recruited class, as well as collecting data on the whole class. Effect sizes for the impact of the game on individual students across these two studies were moderate to large. In the fourth-class group, effects for one target student, Ben, were similar to the effects on the whole class, however effects for another target student, Adam, were not as large as for the whole class. In the senior infants class, the CBGG was not as effective in targeting individual behaviour as it was in targeting whole class behaviour. This demonstrates that although an intervention may appear effective across a whole class group, each child responds differently, and some will respond more favourably than others. In Chapter 5, environmental variables out of the control of the researcher appeared to impact individual student behaviour, particularly Katie’s AEB and DB. Changes in the seating plan outlined in Chapter 5 led to changes in Katie’s behaviour, making it difficult to assert how much of the changes in behaviour could be attributed to the CBGG. The apparent impact of the change in seat on Katie’s behaviour could not have been detected had data not been collected on her behaviour independently of the whole class. Overall, all four individually monitored children responded positively to the CBGG generally and none could be classed as complete ‘non-responders’.

Groves and Austin (2020) evaluated the CBGG with target students versus non-target peers (i.e., students with low levels of disruptive behaviour at baseline). In their study, the target students demonstrated much larger changes in behaviour than did their non-target peers. Importantly however, the non-target peers demonstrated much lower rates of disruption at baseline, making large improvements unlikely. In the evaluations conducted in Chapters 4 and 5 of the current thesis, the behaviour of target students was compared to behaviour of the rest of the class collectively rather than to non-target peers. Both approaches to examining the behaviour of target children are valid, however the difference between the studies in Chapters 4 and 5 and the study by Groves and Austin (2020) means that a direct comparison is not possible. In both fourth class (Chapter 4) and senior infants (Chapter 5), the 'whole class data' reflected data on a diverse mix of students. This included some who were potentially as disruptive as target students, and some who displayed very low rates of disruption. In some cases, the collective behaviour of the whole class at baseline was similar to that of target students. For example, in Chapter 4, whole class AEB occurred during an average of 76.33% of intervals during baseline, and Adam was engaged during a mean of 76.17% of intervals during baseline. In Chapter 5, Ellie's DB (M = 25%) was very similar to that of the whole class (M = 27.5%) at baseline. In most cases, the whole class responded more stably to the intervention than did target students. Target students were more prone to general instability in data collected on their behaviour and their behaviour tended to fluctuate at times based on other environmental changes (e.g., Katie's seat changes in Chapter 5). It is also important to note that although only two target students were chosen for individual monitoring, both teachers stated that they could have chosen more. Final decisions on who to monitor were based on students' attendance, whether they returned consent forms and teacher perceptions about the severity of their behaviour. Going forward, researchers may consider what is the most appropriate way to collect data in classrooms where whole-class interventions will be applied; is it most appropriate to collect data on targets as well as the whole class, targets only (e.g., Tanol et al., 2010), or perhaps targets as well as non-target peers (e.g., Groves & Austin, 2020)? The current findings suggest that in a class where disruption is a wide issue, with many students engaging in disruptive behaviours, whole class monitoring is useful alongside target student monitoring. In cases where there are no behaviour issues outside of target students, perhaps time is best spent collecting data on target students only.

Treatment Integrity. When considering the effectiveness of the CBGG across the diverse populations in this thesis, it is important to consider the treatment integrity across studies as an important contextual factor. Treatment integrity was monitored across all four studies and five classrooms and could range from very low on occasion (e.g., 27.27% in Study 2, Chapter 3) to 100% across all teachers. The fact that the CBGG maintained general effectiveness across classrooms despite low treatment integrity during some sessions suggests that teachers do not need to strive for perfection every time they play the game. It does however point to the need for further research which aims to delineate the steps of the game which are absolutely necessary for its success. Incidents of low treatment integrity were addressed in this thesis by bringing them to the attention of the teacher in person or via email, however future research may consider liaising with the teacher at the outset of the study to discuss the preferred means of addressing this issue. For example, an email prompt may be a useful strategy, however a teacher may not read their emails until the end of the school day and therefore not receive the prompt before class. Some form of self-monitoring of treatment integrity may be a useful avenue for future research.

Evaluating Procedural Variations of the CBGG

There have been recent calls to examine procedural variations during the GBG (Joslyn et al., 2019) and recent research has made strides in identifying components of the game which can be varied without compromising effectiveness. These flexible features have included the awarding of points rather than fouls (i.e., implementing the CBGG rather than the GBG) and playing the game with independent rather than interdependent group contingencies in place (Groves & Austin, 2017). Another recent study found that the GBG could be effective with the whole class playing on one team (Ford et al., 2020). Other studies have evaluated different methods of awarding points during the CBGG by incorporating Class Dojo (Ford, 2017; Lynne et al., 2017) and another earlier mentioned study evaluated the CBGG with a known versus unknown criterion (Groves & Austin, 2020). Bowman-Perrott et al. (2016) considered modified GBG applications as a moderator in their review of GBG research. Despite noting that 11 included studies evaluated a modified GBG and that modified versions of the GBG produced very slightly larger effects than the original GBG, the authors did not state what constituted a ‘modification’. The current thesis systematically evaluated two distinct features of the CBGG: the importance of overt visual feedback and the thinning of the schedule of reinforcement positive behaviour. Although not systematically evaluated, two studies also looked at whether the CBGG was effective with weekly prizes only (Chapter 3).

Teacher buy-in is crucial when developing intervention-based studies. Through demonstrating that features of game-based interventions like the CBGG may be flexible, it may be possible to give teachers a choice in how they choose to implement the game (Groves & Austin, 2020). As the intervention agent, it is important that the teacher maintains a sense of agency and autonomy to maximise their likelihood of adopting a research-based approach (Joram et al., 2020). Time pressure is a commonly cited stressor in the teaching profession (ASTI, 2018), and is a stressor contributing to teacher burnout, particularly emotional exhaustion (Skaalvik & Skaalvik, 2017). The deleterious effects of stress and burnout in the teaching profession are well-established: decreased job satisfaction (Yorulmaz et al., 2017), turnover intentions (Rajendaran et al., 2020) and poorer student outcomes (i.e., behaviour and academic achievement; Herman et al., 2018). Intervention implementation can be time-consuming, however, so too can dealing with disruptive student behaviour. If a simplified version of the CBGG can be put in place with significant time-saving measures included, this could enhance teacher perceptions of the intervention. Several time-saving measures have been supported in this thesis, for example delayed feedback delivery and weekly prizes (with adolescent populations) and schedule thinning. These will be discussed in detail next.

Visual Feedback. Visual feedback has been examined in the context of the GBG previously however, many of these evaluations are now dated (e.g., Harris & Sherman, 1973; Medland & Stachnik, 1972) and more recent evaluations took place with very young children in preschool classrooms (Foley et al., 2019; Wiskow et al., 2019). In the current thesis, visual feedback was examined in the form of visible points drawn onto a scoreboard in Chapter 3. The CBGG-d (a version of the CBGG where visual feedback was withheld until the end of the game) was compared with the CBGG-i (a version where visual feedback was provided on an ongoing basis during the game). Participants were in first year in secondary school. Results indicated that the CBGG was effective both with and without immediate visual feedback in the form of points, however the effects differed across different classroom settings. In Ms. Allen's class, the CBGG-d was more effective than the CBGG-i in targeting AEB, in Ms. Brady's class, the CBGG-i was slightly more effective, whereas in Mr. Carroll's class, the differences in effectiveness in targeting AEB were minimal. The CBGG-d and CBGG-i had similar large effects in targeting DB in Ms. Allen's class, the CBGG-i was more effective than the CBGG-d in Ms. Brady's class and the CBGG-d was more effective than the CBGG-i in Mr. Carroll's class. Students generally preferred the CBGG-i whereas all teachers preferred the CBGG-d. These results suggest that overt visual feedback is not a completely essential element of the CBGG. Although in one class setting (Ms. Brady's class), the CBGG-i was notably more effective than the CBGG-d, the CBGG-d still had large effects on AEB and moderate effects on DB. Manipulation of feedback as a flexible component during the CBGG still requires further research and the current thesis only evaluated visual feedback. This was considered appropriate given adolescents pre-requisite skills (i.e., they can count points, have an understanding for progress of points and what constitutes more than/less than). Future evaluations may consider vocal feedback and visual + vocal feedback (e.g., Wiskow et al., 2019). Wiskow et al., (2019) found visual + vocal feedback most effective during the GBG in their evaluation with preschoolers, however, note that preschoolers may not have had the skills to understand progression during a visual feedback only phase. As well as this, Wiskow et al. had the experimenter implement the game, which could have meant that student attention was not drawn to points being awarded in the visual only condition. This reiterates the importance of considering the population under study when examining varying components during the game. Some variations may be more suitable for use with some populations (e.g, older populations) than others.

The importance of time-saving measures in second-level settings has been outlined earlier and the potential for visual feedback to be withheld until the end of the game has important time-saving implications. This may be particularly relevant to secondary school teachers, who may only see their class for 35-40 min per day. The central issue with having to provide immediate feedback during the CBGG, is that the teacher may be anywhere in the room and must make the effort to approach the board and record the team scores. By withholding feedback, the teacher can privately administer points from wherever he/she is in the room, which has the potential to maximise instruction time. As noted earlier, this flexible feature may be left to the teacher to decide on whether to use it or not, enhancing their scope for autonomous delivery of the intervention. Teachers in Chapter 3 preferred the CBGG-d to the CBGG-i, however this only reflects three teachers' opinions. Further research is clearly warranted on the provision of visual feedback during the CBGG, with all age groups and class levels.

Weekly Prizes. An unexpected ancillary finding emerging from this research was the effectiveness of the CBGG being sustained when prizes were offered weekly only rather than daily or daily *and* weekly with adolescent students (Chapter 3). This procedural variation was incorporated into the studies in Chapter 3 based on teacher feedback during training and is a time-saving mechanism which, when combined with other mechanisms like withholding visual feedback, may make the CBGG accessible for secondary school teachers. The implications of this finding were briefly discussed in Chapter 3, however in examining this finding in a broader context, its true value can be appreciated. Saving 3-5 min per day on the provision of prizes has the potential to add up to a whole class-worth of time over 7-12 days of CBGG implementation. It is also a more economical approach with less resources required through the provision of prizes. This is particularly important in at-risk school settings such as those recruited in Chapter 3. Weekly prizes were examined here as a matter of necessity given teacher concerns and it therefore was not possible to directly compare conditions where weekly prizes were available with conditions where daily or daily *and* weekly prizes were available. Given the potential for the version where prizes are administered weekly and the large amount of time saved this is a hugely important avenue for future evaluations with this age cohort. Weekly prizes only were not evaluated with the primary school children taking part in Chapters 4 and 5, as teachers did not express any concern about administering prizes daily. Primary school teachers' daily routine differs significantly from that of secondary school teachers, in that they teach the same cohort of students in the same room all day. This may leave more time for intervention elements such as awarding of small, daily prizes. Overall, the provision of weekly prizes only in Chapter 3 has important implications for practice in secondary school settings, introducing a time and cost-effective strategy which may be used during the CBGG.

It is evident that there are many different methods of administering prizes during game-based interventions which warrant systematic evaluation. Incorporating 'mystery' elements into game-based classroom management has been a feature in recent research. For example, Kelshaw-Levering et al. (2000) randomised several components of a group contingency game successfully, with students only finding out the target behaviour, the criterion, the type of group contingency (i.e., dependent or interdependent) and reward at the end of the game. The game was effective in producing reductions in disruptive behaviour across the second-grade class with which it was implemented. Robichaux and Gresham (2014) implemented the 'Mystery Motivator' intervention, a game-based intervention whereby students sometimes received a reward for meeting the game criterion, but sometimes did not. At the end of the game, the teacher would determine if the class met a daily goal, and if they did, she would colour the day on a calendar chart. The letter 'M' was hidden in a magic marker on some days on the chart, but not others. The game produced large effects on disruptive behaviour across three classes (see review in Chapter

2). The method applied by Robichaux and Gresham differs from the method applied in Chapter 3 in that students had the potential to earn prizes daily and would have behaved on the premise the prizes may be available. The adolescents taking part in Chapter 3 knew prizes would not be available until the end of a CBGG phase (i.e., the end of a week). As well as comparing weekly prizes only with daily and weekly or daily prizes only (as recommended in the previous paragraph), future research may also consider comparing known weekly prizes with an unknown mystery motivator. The preliminary findings presented here, indicating that weekly prizes *only* are potentially sufficient when implementing the game with adolescents, is a promising development and certainly points towards the usefulness of further future examination of this feature.

Schedule Thinning. Schedules are a common feature in game-play generally (Linehan et al., 2015) and play a role in the implementation of the GBG and CBGG. Despite the central role of schedules of reinforcement in the implementation of these game-based interventions, they had not yet been evaluated as a procedural variation. Chapter 5 addressed this gap in the literature by examining schedule thinning during the CBGG with a senior infants class. Students were first exposed to the CBGG with 2 min intervals between opportunities to earn points. Once the effectiveness of the game had been established, the schedule was thinned across three phases: with 3 min, 4 min and 5 min intervals between opportunities to earn points. CBGG effectiveness was largely sustained across these three phases, with AEB remaining high and DB remaining low. The teacher also liked the thinning element, stating it was *"more manageabl*e" when there was 5 min between behaviour checks. Schedule thinning had been incorporated in recent research on the CW-FIT (e.g., Conklin et al., 2017; Naylor et al., 2018), however the authors did not clearly outline how and when the thinning took place. Chapter 5 therefore provides a first example of schedule thinning during the CBGG, by systematically thinning the schedule of reinforcement over several days.

Schedule thinning during the CBGG has important implications for practice, however, does warrant further research. With further evaluations on thinning during the CBGG with a wider variety of populations, teachers may opt to implement the CBGG with short intervals between points initially, and quickly begin to thin the schedule when DB has reached an acceptable level. Future research should also examine initial interval lengths. Chapter 5 began with 2 min intervals based on previous work on the CBGG and on what the teacher deemed feasible to implement. Future research may begin with larger intervals. Indeed, the studies in Chapter 3 implemented a thin schedule by only having teachers provide reinforcement approximately every 5 min on a variable schedule. This decision was made in conjunction with the teachers involved, however it may have been more beneficial to begin with a denser schedule and thin it over time. Future research may consider comparing implementation of the CBGG with an initial dense schedule of reinforcement and thinning the schedule over time, to implementation of the CBGG with a thin schedule from the outset (e.g., 5 min). Another point for future research which was raised in Chapter 5, is that schedule thinning may be an option for fading intervention intensity within game sessions, for example, beginning the game with points available every 2 min and fading the intensity of the intervention during the same game session by slowly increasing intervals between points. It may also be appropriate to determine the interval length based on interresponse time of DB of the most disruptive students in the class. As mentioned earlier, it is inherently important to consider the population and context under study when evaluating procedural variations during the CBGG as there may be variations which are more appropriate in some contexts than others. For example, if the secondary school teachers in Chapter 3 had been asked to maintain a very dense schedule of reinforcement, this may have decreased their buy-in given their evident time concerns. Despite the need for future research, the current findings, paired with the evaluation of the other procedural variations, provide further evidence for allowing some element of teacher autonomy during CBGG implementation.

The Effects of the CBGG on Teacher Behaviour

Teacher behaviour was not a main focus of this piece of research, however Chapter 3 Study 2 did evaluate teacher reprimands and praise statements during the CBGG. To reiterate the findings presented in that study, the CBGG did not appear to impact reprimand or praise statements by the teachers and in general, teacher praise statements were very low across study phases. These findings were similar to those by Wahl et al., (2016). Data on teacher statements were not collected in subsequent studies due to time constraints in data collection. There was not enough time in a 20 min period to collect enough data on the whole class and target student behaviour as well as teacher behaviour. Student behaviour was a priority in this thesis, so it was decided not to monitor teacher behaviour further. There is the potential that the CBGG would have had a differing impact on primary school teachers' rate of praise and/or reprimands. The implications of collecting data on teacher praise and reprimands and the potential for further research were outlined in Chapter 3. To reiterate, praise has been found to impact AEB and DB of adolescent students when applied alone, without any other intervention (Blaze et al., 2014), and therefore it may have the potential to elevate intervention effects if manipulated during the CBGG. It has also been found however, that adolescents tend to have mixed attitudes towards praise (Elwell & Tiberio, 1994; Fefer et al., 2016). This again points towards the importance of consideration of class context and preferences when deciding to implement a classroom management intervention.

Social Validity of Game-based Interventions

Game-based classroom management has received largely positive ratings from teachers and students in this research. The review chapter revealed that where social validity data was collected from students and teachers, it was largely positive. In Chapters 3-5, social validity data from teachers and students rating the CBGG was predominantly positive, with teachers signifying some important areas which may improve social validity. Time constraints was one issue which may have impacted ratings by some teachers (e.g., Ms. Brady in Chapter 3 referring to being '*overstretched*'), and another teacher noted how the game was effective and acceptable, however when the game ended, students tended to revert to their old ways, engaging in rule-breaking behaviour (Ms. Daly in Chapter 4). These social validity assessments are crucially important for determining avenues for future research on the GBG, CBGG and other game-based classroom management interventions. There is the potential that teachers could have foreseen some of their social validity concerns prior to intervention application but they were not voiced because no formal measure of perceived social validity was conducted at that stage. A pre-intervention acceptability measure may be useful in future research to potentially address teacher concerns before they arise during intervention implementation.

Collection of social validity data from students is also important. A game-based intervention is not worth implementing if students rate the procedure unfavourably, particularly given the positive motivational connotations usually associated with a 'game' or gamification. Students across the three studies for which data were collected on social validity (Chapter 3; Chapter 5), rated the CBGG predominantly positively. This aligns with the findings in Chapter 2, whereby all studies which reported on student social validity, reported positive ratings. Game-based interventions are clearly enjoyed by students, however perhaps more detailed accounts of social validity could be acquired in future research. For example, focus groups where students can articulate their feelings on an intervention may be suitable, particularly for older students taking part in intervention studies.

Game-based Classroom Management

Much of the current chapter thus far has dealt with the specific results included in the empirical thesis chapters and the implications of these results when taken together. Considering the results from a wider lens, findings have implications for the broader field of game-based classroom management research. Although the thesis did not explicitly deal with the benefits of packaging classroom management interventions as games, some benefits are clear. For example, they may provide a teacher with an accessible way to apply well-established behavioural principles in the classroom and they may encourage otherwise disengaged students to engage with classroom management processes. Implications of the findings in this thesis for the field of game-based classroom management research will be considered now. These implications are around terminology used in the discussion on games and cultural contexts in which they are applied.

Game-based Interventions or Gamification?

Earlier in the thesis (Chapters 1 & 2), the concept of gamification was explored as being potentially important in the discussion of game-based classroom management interventions. Its definition, “the use of game design elements in non-game contexts" (Deterding, Dixon et al., 2011, p.10) appeared to fit well with the type of behavioural interventions under study and recent discourse has looked at how gamification and ABA are potentially linked (e.g., Linehan et al., 2015; Morford et al., 2014). Discussion linking the GBG and gamification are, however, rare. Gamification was included as a search term in the systematic review but subsequently was not found to feature in any of the included studies. Throughout the remainder of the thesis, it was therefore not adopted as a term to describe the intervention under study; the CBGG. Although the term gamification was not adopted here and has not been adopted by researchers working in game-based classroom management previously, this should not preclude further discussion on the potential suitability of the term. Misconceptions around gamification, such as the misconception that it must incorporate a digital element, cause potential issues in accepting it as a descriptive term. Deterding and colleagues may even contribute to the generation of these misconceptions. Specifically, they described gamification as “video game elements in non-gaming systems" in one publication (Deterding, Sicart et al., 2011, p. 2425), but gamification is referred to as “the use of game design elements in non-game contexts” in another publication (Deterding, Dixon et al., 2011, p.10). Although the latter constitutes the official proposed definition, lack of consistency may be contributing to certain misconceptions. In Morford et al’s (2014) discussion on gamification and behaviour analysis, there is a heavy focus on video games throughout the article, with many of the examples of gamification given being digital examples (e.g., Zombies Run!, PowerHouse). As outlined in Chapter 1, it is not within the scope of this thesis to discuss the relationship between gamification and ABA in detail, however, the thesis supports the need for such a conversation, in that it may enhance the language at our disposal for discussion on game-based interventions.

Game-based Interventions in Ireland and Europe

The current thesis has served as an important first step in conducting this type of research in an Irish, and even wider European context. As evidenced in Chapter 2, a large majority of the high-quality research on game-based classroom management has been conducted in the USA, with only one included study being conducted outside of the USA (Nolan et al., 2014; conducted in Belize, Central America). Bowman-Perrott et al., (2016) had previously noted that much of the research on the GBG had been conducted in the USA, however that the research tended to cover a range of ethnic groups within US schools. The review in Chapter 2 signalled the need for research outside of the USA on game-based interventions in classroom management, however, the movement appears to have gained some traction since the review was conducted. Apart from the research presented here, Groves and Austin (2020) have conducted important research on the CBGG with Welsh mainstream school populations. This signifies that researchers are realising the potential for these strategies in a European context and that research has and will hopefully continue to build in this area. The findings from the review in Chapter 2 should further spur researchers to conduct investigations in this important area. This thesis has gone some way towards evaluation of game-based interventions in a European, and more specifically an Irish context, however there are many other games which have yet to be examined in Ireland (e.g., the Quiet Classroom Game; Radley et al., 2016), and there are many countries worldwide which are lacking in published literature on game-based interventions. Even the popular GBG has not been applied widely outside of the USA.

Implications for Future Research and Practice

Implications for future research on game-based classroom management interventions have been dispersed throughout this Chapter thus far. One of the first implications related to future research in the field of game-based classroom management is the potential for study on a wide range of game-based interventions, particularly outside of a US context. The review in Chapter 2 identified several unique game-based interventions that have been implemented once only in high-quality research situations. For example, Keep Busy and Carry On (Hine et al., 2015) and the Quiet Classroom Game (Radley et al., 2016), both produced large intervention effects on disruptive and positive classroom behaviours, yet have only been applied once in the literature. There is therefore the potential to evaluate these games with populations across other grade levels and cultures. Another very important point of note from the review chapter, is that virtually all of the game-based interventions in the literature are group contingencies. Therefore, to build upon the findings, a review comparing group contingency interventions framed as games versus those not framed as games may be useful to assert whether the ‘game’ label is important.

The focus of the current thesis was the CBGG, based on its already established effectiveness and the potential for research with Irish populations, particularly those deemed ‘at-risk’ of educational disadvantage. The research is a strong starting point for research with even more diverse Irish populations. For example, given the prevalence of single-sex education, research on the CBGG could be conducted in single-sex girls’ school settings and single-sex secondary school settings. The three secondary school classes who took part in Chapter 3 were mathematics classes, and mathematics was the ongoing subject during the study in Chapter 5, therefore there is an opening to examine the CBGG in other subject settings. There is also scope for research with older secondary school students, such as those preparing for state examinations.

Within CBGG research, there is great scope to investigate distinct procedural variations. The findings related to procedural variations implemented in the current thesis have already been discussed in this chapter, and while promising, there is much more to learn about varying components of the game. Putting mechanisms in place at the end of a game to sustain positive behaviour is an important avenue for future research, as well as further study on feedback necessity and weekly prizes. Further study is also needed about the provision of praise during the CBGG. As discussed earlier, praise may elevate the intervention effects beyond the already established effectiveness, but some types of praise, for example loud, public praise, may not be desired by students (Elwell & Tiberio, 1994).

Strengths and Limitations of the Programme of Research

Strengths

The current programme of research maintained a focus on objective research with substantial rigor and thoroughness from the systematic review through to the empirical studies. The systematic review incorporated reviewer agreement checks during all phases from abstract screening to data extraction and study coding. This rigor was continued throughout the empirical studies in Chapters 3-5, whereby IOA observations ensured that data collection was conducted with objectivity and precision. Efforts were made to adhere to the WWC design standards (WWC, 2017; 2020) used to assess study quality in the review in Chapter 2 when conducting the empirical studies and most experiments met the design standards with reservations. At times it was difficult and time-consuming to ensure these high standards were adhered to given the reliance on other individuals (e.g., an observer dropping out of an observation last minute, cancellation of a data collection session by a school, lengthy meetings with the second reviewer to ensure clarity). However, the commitment to ensuring these elements of the research were conducted appropriately greatly enhances the quality of the programme of research.

Another strength of this work pertains to the diverse age ranges covered in the empirical studies. The range of age groups covered in the current programme of research provides an overview of the CBGG applied across mainstream populations. The CBGG had not been tested empirically with Irish mainstream student populations previously and this pioneering piece of work encompasses studies on student populations at three distinct stages: early schooling (senior infants), pre-adolescent (fourth-class) and young adolescents (first-years). Although the research leaves much scope for further work, such as that with older adolescents, it has provided a strong starting point in indicating potential in a range of settings.

The focus on individual cases (both individual students and class groups) throughout this thesis is a significant strength of the piece of work. Application of the CBGG using single-case research designs to evaluate the effects provided the opportunity for in-depth analysis of diverse populations. The scope to alter the intervention in small ways and evaluate individuals within the context of a group were hugely important to the success of this research (Kazdin, 2011). When considering classroom management research, it is inherently important that *the classroom* is at the core of the evaluation. Every classroom setting is different, a diverse combination of children with varying learning histories. As no two groups of children are the same, neither are any two teachers at the helm of a class. Completing these studies using single-case designs provides a rich insight into the effects of a game-based intervention at the class and individual level.

A final significant strength of this piece of research is the consistent focus on positive behaviour management rather than punishment and negative behaviour. This aligns with an international focus on positive behaviour support systems in schools. Teachers are consistently encouraged to draw their attention to positive behaviour where possible and increase their praise to see improvements in behaviour.

Limitations

The primary limitation in this piece of work is that the CBGG was applied across a limited number (*n* = 5) of classrooms. In some studies presented, the recruitment of additional classes would potentially have led to more valuable results. For instance, when examining the CBGG with a single-sex boys' group in Chapter 4, it would have been an interesting comparison to examine the game with a single-sex girls' group and a mixed-sex group of the same age. The studies in Chapter 3 were very similar and had a second classroom been recruited for Study 1, Study 2 may not have been conducted at all. Difficulties in recruitment meant this was not possible. The recruitment difficulties faced here mostly involved non-response from schools, or teachers being concerned about the time commitment involved in part-taking in research. This is perhaps a point of interest for future research which may consider looking at effective recruitment strategies for behavioural interventions. Although conclusions drawn could be stronger with more classrooms, the studies conducted still provide a comprehensive overview of the CBGG in Irish classrooms spanning from senior infants to first year.

Another limitation of the current work was that some studies were not fully completed before school closures as part of the Irish government's response to the Covid-19 pandemic. The studies in Chapters 4 and 5 were ongoing before school closures occurred, and the closures meant that data collection ceased early, leaving no time for the intended schedule thinning in Chapter 4, and cutting data collection during Chapter 5 slightly short also. Student social validity measures were not gathered in Chapter 4 due to early closures, which was unfortunate. Although this was a considerable limitation, it was completely unforeseen and out of researcher control.

Conclusion

Classrooms are fundamental parts of society, the contents of which can have a huge influence on child development and wellbeing. Integral to the successful management of classrooms, is maintaining strategies and interventions to deal with disruptive behaviour, while fostering engagement and prosocial behaviour. In focusing on classroom management, this thesis was centred on a subdivision of the behavioural approach to classroom management and examined game-based interventions. This thesis looked at how the research has evolved to produce modifications of the GBG (namely the CBGG) and novel, unique games. The CBGG was implemented in five Irish classrooms and findings point towards the importance of evaluating game-based interventions across both individuals and groups, in diverse classroom settings across age categories. There are several implications for research and practice and future iterations should maintain a focus on the stakeholders at hand; teachers and students. Ease of access to interventions such as the CBGG should be a priority in future research and continued efforts should be made to sustain positive approaches to behaviour management in school settings. This programme of research has provided valuable insight into the effectiveness of various versions of the CBGG in Irish contexts, targeting diverse classrooms.

References

Abramson, J. H. (2011). WINPEPI updated: computer programs for epidemiologists, and their teaching potential. *Epidemiologic Perspectives & Innovations*, *8*(1), 1–9. https://doi.org/https://doi.org/10.1186/1742-5573-8-1

Allen, G. E., Common, E. A., Germer, K. A., Lane, K. L., Buckman, M. M., Oakes, W. P., & Menzies, H. M. (2020). A Systematic Review of the Evidence Base for Active Supervision in Pre-K–12 Settings. *Behavioral Disorders*, *45*(3), 167–182. https://doi.org/10.1177/0198742919837646

Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review*, *12*, 30–44. https://doi.org/10.1016/j.edurev.2014.05.003

ASTI. (2018). *Teachers' Work: Work Demands and Work Intensity, March 2018*. https://www.asti.ie/uploads/media/ASTI_-_Survey_of_teachers_-_March_2018.pdf

Ashworth, E., Humphrey, N., & Hennessey, A. (2020). Game over? No main or subgroup effects of the Good Behavior Game in a randomized trial in English primary schools. *Journal of Research on Educational Effectiveness*, *13*(2), 298-321. https://doi.org/10.1080/19345747.2019.1689592

Ayllon, T., & Roberts, M. D. (1974). Eliminating Discipline Problems by Strengthening Academic Performance. *Journal of Applied Behavior Analysis*, *7*(1), 71–76. https://doi.org/10.1901/jaba.1974.7-71

Azoubel, M. S., & Pergher, N. K. (2017). Levantamento sobre a utilização de jogos na Análise do Comportamento Aplicada (A survey on the use of games in Applied Behavior Analysis). *Perspectivas Em Análise Do Comportamento*, *8*(2), 215–225. https://doi.org/10.18761/PAC.2016.014

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some Current Dimensions of Applied Behavior Analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91–97. https://doi.org/10.1901/jaba.1968.1-91

Bagley, W. (1907). *Classroom management*. Macmillan.

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

Barth, J. M., Dunlap, S. T., Dane, H., Lochman, J. E., & Wells, K. C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology*, *42*, 115–133. https://doi.org/10.1016/j.jsp.2003.11.004

Barton, E. E., Pustejovsky, J. E., Maggin, D. M., & Reichow, B. (2017). Technology-aided instruction and intervention for students with ASD: A meta-analysis using novel methods of estimating effect sizes for single-case research. *Remedial and Special Education*, *38*(6), 371–386. https://doi.org/10.1177/0741932517729508

Beeks, A., & Graves Jr., S. (2016). The Effects of the Mystery Motivator Intervention in an Urban Classroom. *School Psychology Forum*, *10*(2), 142–156.

Bergstrom, R., Tarbox, J., & Gutshall, K. A. (2011). Behavioral intervention for domestic pet mistreatment in a young child with autism. *Research in Autism Spectrum Disorders*, *5*(1), 218–221. https://doi.org/10.1016/j.rasd.2010.04.002

Blaze, J. T., Olmi, D. J., Mercer, S. H., Dufrene, B. A., & Tingstom, D. H. (2014). Loud versus quiet praise: A direct behavioral comparison in secondary classrooms. *Journal of School Psychology*, *52*, 349–360. https://doi.org/10.1016/j.jsp.2014.05.006

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis* (1st ed.). John Wiley & Sons, Ltd.

Bostow, D., & Geiger, O. G. (1976). Good Behavior Game : A Replication and Systematic Analysis with a Second Grade Class. *School Applications of Learning Theory of Learning Theory*, *8*(2), 18–27.

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

Brophy, J. (2006). History of Research on Classroom Management. In C. M. Everston & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 17–43). Lawrence Erlbaum Associates, Inc.

Bushell Jr., D., Wrobel, P. A., & Michaelis, M. L. (1968). Applying "Group" Contingencies to the Classroom Study Behavior of Preschool Children. *Journal of Applied Behavior Analysis*, *1*(1), 55–61. https://doi.org/10.1901/jaba.1968.1-55

Casados, M. C. (2012). *Transporting an Evidence-Based Behavioral Intervention to Practice: Treatment Outcomes and Barriers to Implementation* [Doctoral dissertation, The University of Wisconsin-Madison]. https://search.proquest.com/openview/0486f25bf748d1bb405109b884ca67de/1?pq-origsite=gscholar&cbl=18750&diss=y

Central Statistics Office. (2020). *Mainstream National Schools and Pupils in Ordinary Classes by Sex Category School, Year and Statistic - StatBank - data and statistics*. https://statbank.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=EDA51&PLanguage=0

Chaffee, R. K., Briesch, A. M., Johnson, A. H., & Volpe, R. J. (2017). A Meta-Analysis of Class-Wide Interventions for Supporting Student Behavior. *School Psychology Review*, *46*(2), 149–164. https://doi.org/10.17105/SPR-2017-0015.V46-2

ClassDojo. (2019). https://www.classdojo.com/en-gb/?redirect=true

Cochran, W. G. (1954). The Combination of Estimates from Different Experiments. *Biometrics*, *10*(1), 101–129.

Collier-Meek, M. A., Fallon, L. M., & DeFouw, E. R. (2017). Toward Feasible Implementation Support: E-Mailed Prompts to Promote Teachers' Treatment Integrity. *School Psychology Review*, *46*(4), 379–394. https://doi.org/10.17105/SPR-2017-0028.V46-4

Common, E. A., Lane, K. L., Cantwell, E. D., Brunsting, N. C., Oakes, W. P., Germer, K. A., & Bross, L. A. (2020). Teacher-Delivered Strategies to Increase Students' Opportunities to Respond: A Systematic Methodological Review. *Behavioral Disorders*, *45*(2), 67–84. https://doi.org/10.1177/0198742919837646

Conklin, C. G., Kamps, D., & Wills, H. (2017). The Effects of Class-Wide Function-Related Intervention Teams (CW-FIT) on Students' Prosocial Classroom Behaviors. *Journal of Behavioral Education*, *26*(1), 75–100. https://doi.org/10.1007/s10864-016-9252-5

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Pearson/Merrill-Prentice Hall.

Dadakhodjaeva, K. (2017). *The Good Behavior Game: Effects On and Maintenance of Behavior in Middle-School Classrooms Using Class Dojo* [Doctoral dissertation, The University of Southern Mississippi]. https://search.proquest.com/docview/1797920891?accountid=15272

Dadakhodjaeva, K., Radley, K. C., Tingstrom, D. H., Dufrene, B. A., & Dart, E. H. (2019). Effects of Daily and Reduced Frequency Implementation of the Good Behavior Game in Kindergarten Classrooms. *Behavior Modification*, 1–25. https://doi.org/10.1177/0145445519826528

Darch, C. B., & Thorpe, H. W. (1977). The Principal Game: A Group Consequence Procedure to Increase Classroom On-Task Behavior. *Psychology in the Schools*, *14*(3), 341–347. https://doi.org/10.1002/1520-6807(197707)14:3<341::AID-PITS2310140315>3.0.CO;2-3

Darmody, M., Smyth, E., & McCoy, S. (2008). Acting up or opting out? Truancy in Irish secondary schools. *Educational Review*, *60*(4), 359-373. https://doi.org/10.1080/00131910802393399

Davies, S., & Witte, R. (2000). Self-Management and Peer-Monitoring within a Group Contingency to Decrease Uncontrolled Verbalizations of Children with Attention-Deficit/Hyperactivity Disorder. *Psychology in the Schools*, *37*(2), 135–147. https://doi.org/10.1002/(SICI)1520-6807(200003)37:2<135::AID-PITS5>3.0.CO;2-U

De Byl, P. (2013). Factors at Play in Tertiary Curriculum Gamification. *International Journal of Game-Based Learning*, *3*(2), 1–21. https://doi.org/10.4018/ijgbl.2013040101

Dejager, B., Houlihan, D., Filter, K. J., Mackie, P. F. E., & Klein, L. (2020). Comparing the Effectiveness and Ease of Implementation of Token Economy, Response Cost, and a Combination Condition in Rural Elementary School Classrooms. *Journal of Rural Mental Health*, *44*(1), 39–50. https://doi.org/10.1037/rmh0000123

Department of Education and Skills. (2020). Key Statistics 2018/2019 and 2019/2020. In *Department of Education and Skills*.

Department of Education, (1982). The Abolition of Corporal Punishment in National Schools. https://www.education.ie/en/Circulars-and-Forms/Active-Circulars/p_abolition_corporal_punishment.pdf

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification." *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification using game-design elements in non-gaming contexts. *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '11*. https://doi.org/10.1145/1979742.1979575

Dillenburger, K., & Coyle, C. (2019). Education for all: The good inclusion game. *Behavioral Interventions*, *34*, 338–351. https://doi.org/10.1002/bin.1671

Dillon, M. B. M., Radley, K. C., Tingstrom, D. H., Dart, E. H., Barry, C. T., & Codding, R. (2019). The Effects of Tootling via ClassDojo on Student Behavior in Elementary Classrooms. *School Psychology Review*, *48*(1), 18–30. https://doi.org/10.17105/SPR-2017-0090.V48-1

Domitrovich, C. E., Pas, E. T., Bradshaw, C. P., Becker, K. D., Keperling, J. P., Embry, D. D., & Ialongo, N. (2015). Individual and School Organizational Factors that Influence Implementation of the PAX Good Behavior Game Intervention. *Prevention Science*, *16*, 1064–1074. https://doi.org/10.1007/s11121-015-0557-8

Donaldson, J. M., Matter, A. L., & Wiskow, K. M. (2018). Feasibility of and Teacher Preference for Student-Led Implementation of the Good Behavior Game in Early Elementary Classrooms. *Journal of Applied Behavior Analysis*, *51*(1), 118–129. https://doi.org/10.1002/jaba.432

Donaldson, J. M., Vollmer, T. R., Krous, T., Downs, S., & Berard, K. P. (2011). An Evaluation of the Good Behavior Game in Kindergarten Classrooms. *Journal of Applied Behavior Analysis*, *44*(3), 605–609. https://doi.org/10.1901/jaba.2011.44-605

Donaldson, J. M., Wiskow, K. M., & Soto, P. L. (2015). Immediate and distal effects of the good behavior game. *Journal of Applied Behavior Analysis*, *48*(3), 685–689. https://doi.org/10.1002/jaba.229

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

Drevon, D., Fursa, S. R., & Malcolm, A. L. (2017). Intercoder Reliability and Validity of WebPlotDigitizer in Extracting Graphed Data. *Behavior Modification*, *41*(2), 323–339. https://doi.org/10.1177/0145445516673998

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*, 43–51. https://doi.org/10.1016/0022-4405(91)90014-I

Elwell, W. C., & Tiberio, J. (1994). Teacher praise: What students want. *Journal of Instructional Psychology*, *21*(4), 322.

Emmer, E. T., & Stough, L. M. (2001). Classroom Management: A Critical Part of Educational Psychology, With Implications for Teacher Education. *Educational Psychologist*, *36*(2), 103–112. https://doi.org/10.1207/S15326985EP3602_5

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

Fefer, S., DeMagistris, J., & Shuttleton, C. (2016). Assessing adolescent praise and reward preferences for academic behavior. *Translational Issues in Psychological Science, 2*(2), 153–162. https://doi.org/10.1037/tps0000072

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of Reinforcement*. Prentice-Hall, Inc.

Fishbein, J. E., & Wasik, B. H. (1981). Effect of the Good Behavior Game on Disruptive Library Behavior. *Journal of Applied Behavior Analysis*, *14*(1), 89–93. https://doi.org/10.1901/jaba.1981.14-89

Fitbit Charge 2TM. (2020). https://www.fitbit.com/ie/shop/charge2

Floress, M. T., Rock, A. L., & Hailemariam, A. (2017). The Caterpillar Game: A Classroom Management System. *Psychology in the Schools*, *54*(4), 385–403. https://doi.org/10.1002/pits.22000

Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega Jr., R. (2014). Effects of the Good Behavior Game on Challenging Behaviors in School Settings. *Review of Educational Research*, *84*(4), 546–571. https://doi.org/10.3102/0034654314536781

Flower, A., McKenna, J., Muething, C. S., Pedrotty Bryant, D., & Bryant, B. R. (2014). Effects of the Good Behavior Game on Classwide Off-Task Behavior in a High School Basic Algebra Resource Classroom. *Behavior Modification*, *38*(1), 45–68. https://doi.org/10.1177/0145445513507574

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]. https://aquila.usm.edu/dissertations/1046

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

Foxx, R. M., & Rubinoff, A. (1979). Behavioral Treatment of Caffeinism: Reducing Excessive Coffee Drinking. *Journal of Applied Behavior Analysis*, *12*(3), 335–344. https://doi.org/10.1901/jaba.1979.12-335

Gage, N. A., Haydon, T., MacSuga-Gage, A. S., Flowers, E., & Erdy, L. (2020). An Evidence-Based Review and Meta-Analysis of Active Supervision. *Behavioral Disorders*, *45*(2), 117–128. https://doi.org/10.1177/0198742919851021

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, *29*(3), 207–219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Grandy, G. S., Madsen JR., C. H., & de Mersseman, L. M. (1973). The Effects of Individual and Interdependent Contingencies on Inappropriate Classroom Behavior. *Psychology in The Schools*, *10*(4), 488–493. https://doi.org/10.1002/1520-6807(197310)10:4<488::AID-PITS2310100420>3.0.CO;2-F

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

Groves, E. A., & Austin, J. L. (2019). Does the Good Behavior Game evoke negative peer pressure? Analyses in primary and secondary classrooms. Journal of Applied Behavior Analysis, 52(1), 3-16. https://doi.org/10.1002/jaba.513

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

Hall, R. V., Lund, D., & Jackson, D. (1968). Effects of Teacher Attention on Study Behavior. *Journal of Applied Behavior Analysis*, *1*(1), 1–12. https://doi.org/10.1901/jaba.1968.1-1

Hansen, B. D., Caldarella, P., Williams, L., & Wills, H. P. (2017). Managing Student Behavior in Dual Immersion Classrooms: A Study of Class-Wide Function-Related Intervention Teams. *Behavior Modification*, *41*(5), 626–646. https://doi.org/10.1177/0145445517698418

Harris, V. W., & Sherman, J. A. (1973). Use and analysis of the “Good Behavior Game” to reduce disruptive classroom behavior. *Journal of Applied Behavior Analysis*, *6*(3), 405–417. https://doi.org/10.1901/jaba.1973.6-405

Hartman, K., & Gresham, F. (2016). Differential Effectiveness of Interdependent and Dependent Group Contingencies in Reducing Disruptive Classroom Behavior. *Journal of Applied School Psychology*, *32*(1), 1–23. https://doi.org/10.1080/15377903.2015.1056922

Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, *3*(3), 224–239. https://doi.org/10.1002/jrsm.1052

Heekes, S. L., Kruger, C. B., Lester, S. N., & Ward, C. L. (2020). A Systematic Review of Corporal Punishment in Schools: Global Prevalence and Correlates. *Trauma, Violence, and Abuse*, 1–21. https://doi.org/10.1177/1524838020925787

Herman, K. C., Hickmon-Rosa, J., & Reinke, W. M. (2018). Empirically Derived Profiles of Teacher Stress, Burnout, Self-Efficacy, and Coping and Associated Student Outcomes. *Journal of Positive Behavior Interventions*, *20*(2), 90–100. https://doi.org/10.1177/1098300717732066

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

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hine, J. F., Ardoin, S. P., & Foster, T. E. (2015). Decreasing transition times in elementary school classrooms: Using computer-assisted instruction to automate intervention components. *Journal of Applied Behavior Analysis*, *48*(3), 495–510. https://doi.org/10.1002/jaba.233

Hirsch, S. E., Healy, S., Judge, J. P., & Lloyd, J. W. (2016). Effects of an interdependent group contingency on engagement in physical education. *Journal of Applied Behavior Analysis*, *49*(4), 1–5. https://doi.org/10.1002/jaba.328

Hoff, K. E., & Ervin, R. A. (2013). Extending Self-management Strategies: The Use of a Classwide Approach. *Psychology in the Schools*, *50*(2), 151–164. https://doi.org/10.1002/pits.21666

Infantino, J., & Little, E. (2005). Students' Perceptions of Classroom Behaviour Problems and the Effectiveness of Different Disciplinary Methods. *Educational Psychology*, *25*(5), 491–508. https://doi.org/10.1080/01443410500046549

Johnson, D., Deterding, S., Kuhn, K. A., Staneva, A., Stoyanov, S., & Hides, L. (2016). Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions*, *6*, 89–106. https://doi.org/10.1016/j.invent.2016.10.002

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., & Vollmer, T. R. (2020). Efficacy of teacher-implemented Good Behavior Game despite low treatment integrity. *Journal of Applied Behavior Analysis*, *53*(1), 465–474. https://doi.org/10.1002/jaba.614

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

Kamps, D., Conklin, C., & Wills, H. (2015). Use of self-management with the CW-FIT group contingency program. *Education and Treatment of Children*, *38*(1), 1–32. https://doi.org/10.1353/etc.2015.0003

Kamps, D., Wills, H. P., Heitzman-Powell, L., Laylin, J., Szoke, C., Petrillo, T., & Culey, A. (2011). Class-Wide Function-Related Intervention Teams: Effects of Group Contingency Programs in Urban Classrooms. *Journal of Positive Behavior Interventions*, *13*(3), 154–167. https://doi.org/10.1177/1098300711398935

Kazdin, A. E. (2011). *Single-Case Research Designs: Methods for Clinical and Applied Settings* (2nd ed.). Oxford University Press.

Kazdin, A. E., & Bootzin, R. R. (1972). The Token Economy: An Evaluative Review. *Journal of Applied Behavior Analysis, 5*(3), 343–372. https://doi.org/10.1901/jaba.1972.5-343

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

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

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A Meta-Analysis of the Effects of Classroom Management Strategies and Classroom Management Programs on Students' Academic, Behavioral, Emotional and Motivational Outcomes. *Review of Educational Research, 86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Kounin, J. S., & Gump, P. V. (1961). The comparative influence of punitive and nonpunitive teachers upon childrens' concepts of school misconduct. *Journal of Educational Psychology*, *52*(1), 44–49. https://doi.org/10.1037/h0044705

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-Case Design Technical Documentation*. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd.pdf

Landrum, T. J., & Kauffman, J. M. (2006). Behavioral approaches to classroom management. In C. M. Everston & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 47–71). Lawrence Erlbaum Associates, Inc.

Lannie, A. L., & McCurdy, B. L. (2007). Preventing Disruptive Behavior in the Urban Classroom: Effects of the Good Behavior Game on Student and Teacher Behavior.

Education and Treatment of Children, *30*(1), 85–98. https://doi.org/10.1353/etc.2007.0002

Larson, M., Cook, C. R., Fiat, A., & Lyon, A. R. (2018). Stressed Teachers Don't Make Good Implementers: Examining the Interplay Between Stress Reduction and Intervention Fidelity. *School Mental Health*, *10*, 61–76. https://doi.org/10.1007/s12310-018-9250-y

Lastrapes, R. E. (2016). Let Us Play: Using Research-Based Games to Facilitate Effective Instruction. *Beyond Behavior*, *25*(3), 27–33. https://doi.org/10.1177/107429561602500305

LeBlanc, L. A., Hagopian, L. P., Maglieri, K. A., & Poling, A. (2002). Decreasing the intensity of reinforcement-based interventions for reducing behavior: Conceptual issues and a proposed model for clinical practice. *The Behavior Analyst Today*, *3*(3), 289–300. https://doi.org/10.1037/h0099991

Linehan, C., Kirman, B., & Roche, B. (2015). Gamification as Behavioral Psychology. In S. P. Walz & S. Deterding (Eds.), *The Gameful World: Approaches, Issues, Applications*. MIT Press.

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

Long, A. C. J., Miller, F. G., & Upright, J. J. (2019). Classroom management for ethnic–racial minority students: A meta-analysis of single-case design studies. *School Psychology*, *34*(1), 1–13. https://doi.org/10.1037/spq0000305

Lynne, S., Radley, K. C., Dart, E. H., Tingstrom, D. H., Barry, C. T., & Lum, J. D. K. (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., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology*, *49*(5), 529–554. https://doi.org/10.1016/j.jsp.2011.05.001

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

Malecki, C. K., & Elliott, S. N. (2002). Children's Social Behaviors as Predictors of Academic Achievement: A Longitudinal Analysis. *School Psychology Quarterly*, *17*(1), 1–23. https://doi.org/10.1521/scpq.17.1.1.19902

Marsh, T. (2011). Serious games continuum: Between games for purpose and experiential environments for purpose. *Entertainment Computing*, *2*(2), 61–68. https://doi.org/10.1016/j.entcom.2010.12.004

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

McCoy, S., Quail, A., & Smyth, E. (2014). The effects of school social mix: Unpacking the differences. *Irish Educational Studies*, *33*(3), 307-330.

McCoy, S., Smyth, E., & Banks, J. (2012). The primary classroom: Insights from the growing up in Ireland study. *Dublin: ESRI*.

McGoey, K. E., Schneider, D. L., Rezzetano, K. M., Prodan, T., & Tankersley, M. (2010). Classwide Intervention to Manage Disruptive Behavior in the Kindergarten Classroom. *Journal of Applied School Psychology*, *26*(3). https://doi.org/10.1080/15377903.2010.495916

McIntosh, K., Reinke, W. M., Kelm, J. L., & Sadler, C. A. (2013). Gender Differences in Reading Skill and Problem Behavior in Elementary School. *Journal of Positive Behavior Interventions*, *15*(1), 51–60. https://doi.org/10.1177/1098300712459080

McKissick, C., Hawkins, R. O., Lentz, F. E., Hailley, J., & McGuire, S. (2010). Randomizing Multiple Contingency Components to Decrease Disruptive Behaviors and Increase Student Engagement in an Urban Second-grade Classroom. *Psychology in the Schools*, *47*(9), 944–959. https://doi.org/10.1002/pits.21666

Medland, M. B., & Stachnik, T. J. (1972). Good-Behavior Game: A Replication and Systematic Analysis. *Journal of Applied Behavior Analysis*, *5*(1), 45–51. https://doi.org/10.1901/jaba.1972.5-45

Mitchell, C., Harwin, A., Vara-Orta, F., & Sheehan, F. (2017). *Single-Gender Public Schools in 5 Charts - Education Week*. https://www.edweek.org/ew/section/multimedia/single-gender-public-schools-in-5-charts.html#National

Mitchell, M. M., & Bradshaw, C. P. (2013). Examining classroom influences on student perceptions of school climate: The role of classroom management and exclusionary discipline strategies. *Journal of school psychology*, *51*(5), 599-610.

Mitchell, R. R. (2014). *Evaluating The Effectiveness of the Good Behavior Game with General Education High School Students Utilizing a Changing Criterion Component.* University of Southern Mississippi.

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

Mitchell, S. K. (1979). Interobserver Agreement, Reliability, and Generalizability of Data Collected in Observational Studies. *Psychological Bulletin*, *86*(2), 376–390. https://doi.org/10.1037/0033-2909.86.2.376

Mitchem, K. J., Young, K. R., West, R. P., & Benyo, J. (2001). CWPASM: A Classwide Peer-assisted Self-management Program for General Education Classrooms. *Education and Treatment of Children*, *24*(2), 111–140.

Moeyaert, M., Maggin, D., & Verkuilen, J. (2016). Reliability, Validity, and Usability of Data Extraction Programs for Single-Case Research Designs. *Behavior Modification, 40*(6), 874–900. https://doi.org/10.1177/0145445516645763

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7). https://doi.org/10.1371/journal.pmed.1000097

Moore, T. C., Maggin, D. M., Thompson, K. M., Gordon, J. R., Daniels, S., & Lang, L. (2019). Evidence Review for Teacher Praise to Improve Students' Classroom Behavior. *Journal of Positive Behavior Interventions*, *21*(1), 3–18. https://doi.org/10.1177/1098300718766657

Morford, Z. H., Witts, B. N., Killingsworth, K. J., & Alavosius, M. P. (2014). Gamification: The Intersection Between Behavior Analysis and Game Design Technologies. *Behavior Analyst*, *37*, 25–40. https://doi.org/10.1007/s40614-014-0006-1

Mudgal, D. (2006). *Randomizing Criteria within and Interdependent Group Contingency: Effects on Math Seatwork*. [Doctoral disseratation, The University of Southern Mississippi].

National Center for Education Statistics. (2019). *Educational Institutions*. https://nces.ed.gov/fastfacts/display.asp?id=84

National Council for Special Education. (2020). *Mission & Guiding Principles*. https://www.nbss.ie/home/mission-guiding-principles

National Educational Welfare Board. (2008). Developing a Code of Behaviour: Guidelines for Schools. In *Tusla*.

Naylor, A. S., Kamps, D., & Wills, H. (2018). The effects of the CW-FIT group contingency on class-wide and individual behavior in an urban first grade classroom. *Education and Treatment of Children*, *41*(1), 1–30. https://doi.org/10.1353/etc.2018.0000

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

O'Leary, K. D., & Becker, W. C. (1967). Behavior modification of an adjustment class: a token reinforcement program. *Exceptional Children*, *33*(9), 637–642. https://doi.org/10.1177/001440296703300905

O'Leary, K. D., Becker, W. C., Evans, M. B., & Saudargas, R. A. (1969). A Token Reinforcement Program in a Public School: A Replication and Systematic Analysis. *Journal of Applied Behavior Analysis*, *2*(1), 3–13. https://doi.org/10.1901/jaba.1969.2-3

Octopus Watch V2. (2020). https://www.heyjoy.io/collections/frontpage/products/octopus-watch-motion-edition-combo

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Techniques Effect Size in Single-Case Research: A Review of Nine Nonoverlap. *Behavior Modification*, *35*(4), 303–322. https://doi.org/10.1177/0145445511399147

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining Nonoverlap and Trend for Single-Case Research: Tau-U. *Behavior Therapy*, *42*, 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Pas, E. T., Cash, A. H., O 'Brennan, L., Debnam, K. J., & Bradshaw, C. P. (2015). Profiles of classroom behavior in high schools: Associations with teacher behavior management strategies and classroom composition. *Journal of School Psychology*, *53*, 137–148. https://doi.org/10.1016/j.jsp.2014.12.005

Patrick, C. A., Ward, P., & Crouch, D. W. (1998). Effects of Holding Students Accountable for Social Behaviors during Volleyball Games in Elementary Physical

Education. *Journal of Teaching in Physical Education*, *17*, 143–156. https://doi.org/10.1123/jtpe.17.2.143

Pustejovsky, J. E. (2016). *scdhlm: A web-based calculator for between-case standardized mean differences (Version 0.3.1) [Web application]*.

Pustejovsky, J. E., Hedges, L. V., & Shadish, W. R. (2014). Design-Comparable effect sizes in multiple baseline designs: A general modeling framework. *Journal of Educational and Behavioral Statistics*, *39*(5), 368–393. https://doi.org/10.3102/1076998614547577

Quail, A., & Smyth, E. (2014). Multigrade teaching and age composition of the class: The influence on academic and social outcomes among students. *Teaching and Teacher Education*, *43*, 80-90. https://doi.org/10.1016/j.tate.2014.06.004

Radley, K. C., Dart, E. H., & O'Handley, R. D. (2016). The Quiet Classroom Game: A Class-Wide Intervention to Increase Academic Engagement and Reduce Disruptive Behavior. *School Psychology Review*, *45*(1), 93–108. https://doi.org/10.17105/SPR45-1.93-108

Rajendran, N., Watt, H. M. G., & Richardson, P. W. (2020). Teacher burnout and turnover intent. *The Australian Educational Researcher*, *47*(3), 477–500. https://doi.org/10.1007/s13384-019-00371-x

Robertshaw, C. S., & Hiebert, H. D. (1973). The Astronaut Game: A Group Contingency Applied to a First Grade Classroom. *School Applications of Learning Theory*, *6*(1), 28–33.

Robichaux, N. M., & Gresham, F. M. (2014). Differential Effects of the Mystery Motivator Intervention Using Student-Selected and Mystery Rewards. *School Psychology Review*, *43*(3), 286–298. https://doi.org/10.1080/02796015.2014.12087428

Rohatgi, A. (2019). *WebPlotDigitizer* (4.2.). https://apps.automeris.io/wpd/

Saigh, P. A., & Umar, A. M. (1983). The Effects of a Good Behavior Game on the Disruptive Behavior of Sudanese Elementary Students. *Journal of Applied Behavior Analysis*, *16*(3), 339–344. https://doi.org/10.1901/jaba.1983.16-339

Sailer, M., & Sailer, M. (2021). Gamification of in-class activities in flipped classroom lectures. *British Journal of Educational Technology*, *52*(1), 75-90. https://doi.org/10.1111/bjet.12948

Shadish, W. R., Zelinsky, N. A. M., & Vevea, J. L. (2016). A Survey of Publication Practices of Single-case Design Researcher When Treatments Have Small or Large Effects. *Journal of Applied Behavior Analysis*, *49*(3), 656–673. https://doi.org/10.1002/jaba.308

Sharafan, E. (2018). *Tabata Timer: Interval Timer Workout Timer HIIT* (3.6.1). https://play.google.com/store/apps/details?id=com.evgeniysharafan.tabatatimer

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based Practices in Classroom Management: Considerations for Research to Practice. *Education and Treatment of Children*, *31*(3), 351-380. https://doi.org/10.1353/etc.0.0007

Skaalvik, E. M., & Skaalvik, S. (2017). Dimensions of teacher burnout: relations with potential stressors at school. *Social Psychology of Education*, *20*, 775–790. https://doi.org/10.1007/s11218-017-9391-0

Skinner, B. F. (1953). *Science and human behavior.* Macmillan.

Smith, D. J. (2006). *School Experience and Delinquency at Ages 13 to 16*. Edinburgh: University of Edinburgh, Centre for Law and Society.

Smith, S., Barajas, K., Ellis, B., Moore, C., McCauley, S., & Reichow, B. (2019). A meta-analytic review of randomized controlled trials of the good behavior game. *Behavior modification*. https://doi.org/10.1177/0145445519878670

Smyth, E., & Quail, A. (2017). Structuring student behaviour: discipline policy in primary schools. *Irish Teachers' Journal*, *5*(1), 9–24.

Speight, D. R. (2018). *Impact of the Implementation of Class-Wide Function-Related Intervention Teams on Adolescent Student Behavior*. [Doctoral dissertation, University of Arkansas].

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School psychology review*, *35*(2), 245-259. https://doi.org/10.1080/02796015.2006.12087989

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*, 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

Thomas, D. R., Becker, W. C., & Armstrong, M. (1968). Production and Elimination of Disruptive Classroom Behavior by Systematically Varying Teacher's Behavior. *Journal of Applied Behavior Analysis*, *1*(1), 35–45. https://doi.org/10.1901/jaba.1968.1-35

Thorne, S., & Kamps, D. (2008). The effects of a group contingency intervention on academic engagement and problem behavior of at-risk students. *Behavior Analysis in Practice*, *1*(2), 12–18. https://doi.org/10.1007/BF03391723

Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The Good Behavior Game: 1969-2002. *Behavior Modification*, *30*(2), 225–253. https://doi.org/10.1177/0145445503261165

Toussaint, K. A., & Tiger, J. H. (2012). Reducing Covert Self-Injurious Behavior Maintained by Automatic Reinforcement Through a Variable Momentary DRO Procedure. *Journal of Applied Behavior Analysis*, *45*(1), 179–184. https://doi.org/10.1901/jaba.2012.45-179

Valentine, J. C., Tanner-Smith, E. E., Pustejovsky, J. E., & Lau, T. S. (2016). Between-case standardized mean difference effect sizes for single-case designs: a primer and tutorial using the scdhlm web application. *Campbell Systematic Reviews*, *12*(1), 1–31. https://doi.org/10.4073/cmdp.2016.1

van Dijk, W., & Gage, N. A. (2019). The effectiveness of visual activity schedules for individuals with intellectual disabilities: A meta-analysis. *Journal of Intellectual and Developmental Disability*, *44*(4), 384–395. https://doi.org/10.3109/13668250.2018.1431761

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

Vannest, K.J., Parker, R.I., Gonen, O., & Adiguzel, T. (2016). *Single Case Research: web based calculators for SCR analysis. (Version 2.0) [Web-based application]*. College Station, TX: Texas A&M University. Available from singlecaseresearch.org

Vidoni, C., Lee, C. H., & Azevedo, L. B. (2014). Fair Play Game: a group contingency strategy to increase students' active behaviours in physical education. *Early Child Development and Care*, *184*(8), 1127-1141.

Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. https://doi.org/10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the Metafor Package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W. (2020). *Convergence Problems with the rma() Function [The metafor Package]*. http://www.metafor-project.org/doku.php/tips:convergence_problems_rma#example_3convergence_with_adjusted_step_length

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

Walker, H. M. (2004). Commentary: Use of Evidence-Based Interventions in Schools: Where We've Been, Where We Are, and Where We Need to Go. *School Psychology Review*, *33*(3), 398–407. https://doi.org/10.1080/02796015.2004.12086256

Wallace, B. C., Small, K., Brodley, C. E., Lau, J., & Trikalinos, T. A. (2012). Deploying an interactive machine learning system in an evidence-based practice center: abstrackr. *ACM International Health Informatics Symposium (IHI)*, 819–824. http://abstrackr.cebm.brown.edu/citing/

Watkins, L., Ledbetter-Cho, K., O'Reilly, M., Barnard-Brak, L., & Garcia-Grau, P. (2019). Interventions for students with autism in inclusive settings: A best-evidence synthesis and meta-analysis. *Psychological Bulletin*, *145*(5), 490–507. https://doi.org/10.1037/bul0000190

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological review*, *20*(2), 158-177. https://doi.org/10.1037/h0074428

Wehby, J. H., Maggin, D. M., Moore Partin, T. C., & Robertson, R. (2012). The Impact of Working Alliance, Social Validity, and Teacher Burnout on Implementation Fidelity of the Good Behavior Game. *School Mental Health*, *4*, 22–33. https://doi.org/10.1007/s12310-011-9067-4

Weir, S. (2006). A report on the procedures used to identify post-primary schools for inclusion in the School Support Programme under DEIS. *Dublin, Ireland: Education Research Centre.*

What Works Clearinghouse. (2016). *Study Review Guide Instructions for Reviewing Single-Case Designs Studies*. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_srg_scd_instructions_s3_v2.pdf

What Works Clearinghouse. (2017). *What Works Clearinghouse™ Standards Handbook, Version 4.0*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/handbooks

What Works Clearinghouse. (2020). *What Works Clearinghouse™ Standards Handbook, Version 4.1*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/handbooks

Williams, J., Thornton, M., Morgan, M., Quail, A., Smyth, E., Murphy, D., & O'Mahony, D. (2018). *The Lives of 13-Year-Olds*. https://www.growingup.ie/pubs/Lives-of-13-Year-olds-report.pdf

Wills, H. P. (2002). *A Praise Game to Improve Both Teacher and Student Behaviors*. [Doctoral dissertation, University of Kansas].

Wills, H. P., Kamps, D., Hansen, B., Conklin, C., Bellinger, S., Neaderhiser, J., & Nsubuga, B. (2009). The Classwide Function-Based Intervention Team Program. *Preventing School Failure: Alternative Education for Children and Youth*, *54*(3), 164–171. https://doi.org/10.1080/10459880903496230

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* (pp. 251–288).

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/10.1177/1098300711421008

Yorulmaz, Y. I., Colak, I., & Altinkurt, Y. (2017). A Meta-Analysis of the Relationship Between Teachers' Job Satisfaction and Burnout. *Eurasian Journal of Educational Research*, *71*, 175–192. https://doi.org/10.14689/ejer.2017.71.10

Appendices

Appendix A: Search Strategies and Findings Across Databases

1) Education Research Complete

EBSCO Host

Table A-1

Search Strategy Applied and Records Identified Across the Education Research Complete Database

Search	Number of Records
S1 **Game Terms** **Field Code:** All Text (TX) TX (Game* OR gamif* OR "class-wide function related intervention teams" OR cw-fit) OR DE "GAMES"	299, 510
S2 **Behaviour Terms** **Field Codes:** Title, Abstract, Keywords, Subjects TI AB SU KW (Behavio#r OR disrupt* OR conduct OR Management N5 classroom OR discipline OR attendance OR truancy OR absenteeism OR absence OR lateness OR punctuality OR timeliness OR tardiness OR engagement OR preparedness OR	415,076

Search	**Number of Records**
participation OR “active student responding” OR “on-task" OR “off-task” OR “hand-rais*” OR handrais* OR Tootl* OR tattl* OR homework OR “talking-out” OR “talking without permission” OR “seat-leaving” OR “out-of-seat” OR Aggression)	
S3 **Behaviour terms** **Exact Subject Terms** (identified by typing “classroom behavio#r” into thesaurus and using all relevant terms that appeared in list- recommended by librarian) DE "STUDENT engagement" OR DE "CLASSROOMS -- Research" OR DE "CLASSROOM activities" OR DE "CLASSROOM management" OR DE "CLASSROOM management research"	20,295
S4 S2 OR S3	420,848
S5 **School/Grade Terms** **Subject terms 1:** typed ‘kindergarten’ into thesaurus and selected relevant terms returned DE "KINDERGARTEN" OR DE "KINDERGARTEN children"	5273
S6	23,931

Search	**Number of Records**
School/grade terms **Subject terms 2:** typed ‘elementary school” into thesaurus and selected relevant terms returned DE "PRIMARY schools" OR DE "PRIMARY education" OR DE "ELEMENTARY schools"	
S7 **School/grade terms** **Subject terms 3:** typed ‘high school” into thesaurus and selected relevant terms returned DE "HIGH school students" OR DE "JUNIOR high school students" OR DE "SECONDARY schools" OR DE "HIGH schools" OR DE "MIDDLE schools" OR DE "SECONDARY school students"	36,857
S8 S5 OR S6 OR S7	64,565
S9 **Field Codes:** Title, Abstract, Keywords, Subjects Class* OR School* OR kindergarten OR "K-12" OR "Grade 1" OR "Grade 2" OR "Grade 3" OR "Grade 4" OR "Grade 5" OR "Grade 6" OR "Grade 7" OR "Grade 8" OR "Grade 9" OR "Grade 10" OR "Grade 11" OR "Grade 12" OR “first-grade” OR “second-grade” OR “Third-grade” OR “Fourth-grade” OR “fifth-grade” OR “sixth-grade” OR “seventh-grade” OR “eighth-grade” OR “ninth-grade” OR “tenth-grade” OR “eleventh-grade” OR “twelfth-grade” OR “1st grade” OR “2nd grade” OR “3rd grade” OR “4th grade” OR “5th grade” OR “6th grade” OR “7th grade” OR “8th grade” OR “9th grade” OR “10th grade” OR “11th grade” OR “12th grade” OR Mainstream OR “general education” OR “regular education"	982,561
S10 S8 OR S9	984,805

Search	Number of Records
S11 **Design Terms** **Field Code:** TX (all text) TX “single-subject” OR SCD OR SCRD OR “single-case” OR Abab OR reversal OR “reversal design*” OR withdrawal OR “withdrawal design*” OR Baseline N20 treatment OR Baseline N20 intervention OR Baseline N20 game* OR Baseline N20 implement* OR “Alternating treatment*” OR multi*element OR multielement OR “multiple baseline” OR “multiple baseline design” OR “changing criterion” OR “changing criterion design” OR “group contingenc*”	68, 223
Final Search S1 AND S4 AND S10 AND S11	**1, 531**

2) ERIC

Proquest Host

Table A-2

Search Strategy Applied and Records Identified Across the ERIC Database

Search	Number of Records
S1 **Game Terms** **Field Code:** Anywhere (Game OR gamif* OR “class-wide function-related intervention teams” OR cw-fit) OR MAINSUBJECT.EXACT("Games")	24,866
S2 **Behaviour Terms** **Field Codes:** Title, Abstract, subject ti(Behavio*r OR disrupt* OR conduct OR Management NEAR/5 classroom OR discipline OR attendance OR truancy OR absenteeism OR absence OR lateness OR punctuality OR timeliness OR tardiness OR engagement OR preparedness OR participation OR “active student responding” OR “on-task" OR "on task" OR “off-task” OR "off task" OR “hand-rais*” OR handrais* OR Tootl* OR tattl* OR homework OR “talking-out” OR "talking out" OR “talking without permission” OR “seat-leaving” OR "seat leaving" OR “out-of-seat” OR "out of seat" OR Aggression) OR ab(Behavio*r OR disrupt* OR conduct OR Management NEAR/5 classroom OR discipline OR attendance OR truancy OR absenteeism OR absence OR lateness OR punctuality OR timeliness OR tardiness OR engagement OR preparedness OR participation OR “active student responding” OR “on-task" OR "on task" OR “off-task” OR "off task" OR “hand-rais*” OR handrais* OR Tootl* OR tattl* OR homework OR “talking-out” OR "talking out" OR “talking without permission” OR “seat-leaving” OR "seat leaving" OR “out-of-seat” OR "out of seat" OR Aggression) OR su(Behavio*r OR disrupt* OR conduct OR Management NEAR/5 classroom OR discipline OR attendance OR truancy OR absenteeism OR absence OR lateness OR punctuality OR timeliness OR tardiness OR engagement OR preparedness OR participation OR “active student responding” OR “on-task" OR "on task" OR “off-task” OR "off task" OR “hand-rais*” OR handrais* OR Tootl* OR tattl* OR homework OR “talking-out” OR "talking	354,951

Search	Number of Records
out" OR "talking without permission" OR "seat-leaving" OR "seat leaving" OR "out-of-seat" OR "out of seat" OR Aggression)	
S3 **Behaviour terms** **Exact Subject Terms** (identified by typing classroom behavio*r into thesaurus and using all relevant terms that appeared in list-recommended by librarian) MAINSUBJECT.EXACT("Classroom Techniques")	26,646
S4 S2 OR S3	370, 361
S5 **School/Grade Terms** **Subject terms 1:** typed 'kindergarten' into thesaurus and selected relevant terms returned MAINSUBJECT.EXACT("Kindergarten")	10, 869
S6 **School/grade terms** **Subject terms 2:** typed 'elementary school" into thesaurus and selected relevant terms returned MAINSUBJECT.EXACT("Elementary Schools") OR MAINSUBJECT.EXACT("Elementary School Students")	52, 849
S7 **School/grade terms** **Subject terms 3:** typed 'high school" into thesaurus and selected relevant terms returned MAINSUBJECT.EXACT("Junior High School Students") OR MAINSUBJECT.EXACT("High Schools") OR	82,858

Search	Number of Records
MAINSUBJECT.EXACT("Junior High Schools") OR MAINSUBJECT.EXACT("High School Students")	
S8 S5 OR S6 OR S7	141,933
S9 **Field Codes:** Title, Abstract, Subjects ti(Class* OR School* OR kindergarten OR "K-12" OR "Grade 1" OR "Grade 2" OR "Grade 3" OR "Grade 4" OR "Grade 5" OR "Grade 6" OR "Grade 7" OR "Grade 8" OR "Grade 9" OR "Grade 10" OR "Grade 11" OR "Grade 12" OR “first-grade” OR "first grade" OR “second-grade” OR "second grade" OR “Third-grade” OR "third grade" OR “Fourth-grade” OR "fourth grade" OR “fifth-grade” OR "fifth grade" OR “sixth-grade” OR "sixth grade" OR “seventh-grade” OR "seventh grade" OR “eighth-grade” OR "eighth grade" OR “ninth-grade” OR "ninth grade" OR “tenth-grade” OR "tenth grade" OR “eleventh-grade” OR "eleventh grade" OR “twelfth-grade” OR "twelfth grade" OR “1st grade” OR “2nd grade” OR “3rd grade” OR “4th grade” OR “5th grade” OR “6th grade” OR “7th grade” OR “8th grade” OR “9th grade” OR “10th grade” OR “11th grade” OR “12th grade” OR Mainstream OR “general education” OR “regular education") OR ab(Class* OR School* OR kindergarten OR "K-12" OR "Grade 1" OR "Grade 2" OR "Grade 3" OR "Grade 4" OR "Grade 5" OR "Grade 6" OR "Grade 7" OR "Grade 8" OR "Grade 9" OR "Grade 10" OR "Grade 11" OR "Grade 12" OR “first-grade” OR "first grade" OR “second-grade” OR "second grade" OR “Third-grade” OR "third grade" OR “Fourth-grade” OR "fourth grade" OR “fifth-grade” OR "fifth grade" OR “sixth-grade” OR "sixth grade" OR “seventh-grade” OR "seventh grade" OR “eighth-grade” OR "eighth grade" OR “ninth-grade” OR "ninth grade" OR “tenth-grade” OR "tenth grade" OR “eleventh-grade” OR "eleventh grade" OR “twelfth-grade” OR "twelfth grade" OR “1st grade” OR “2nd grade” OR “3rd grade” OR “4th grade” OR “5th grade” OR “6th grade” OR “7th grade” OR “8th grade” OR “9th	848 153

Search	Number of Records
grade" OR "10th grade" OR "11th grade" OR "12th grade" OR Mainstream OR "general education" OR "regular education") OR su(Class* OR School* OR kindergarten OR "K-12" OR "Grade 1" OR "Grade 2" OR "Grade 3" OR "Grade 4" OR "Grade 5" OR "Grade 6" OR "Grade 7" OR "Grade 8" OR "Grade 9" OR "Grade 10" OR "Grade 11" OR "Grade 12" OR "first-grade" OR "first grade" OR "second-grade" OR "second grade" OR "Third-grade" OR "third grade" OR "Fourth-grade" OR "fourth grade" OR "fifth-grade" OR "fifth grade" OR "sixth-grade" OR "sixth grade" OR "seventh-grade" OR "seventh grade" OR "eighth-grade" OR "eighth grade" OR "ninth-grade" OR "ninth grade" OR "tenth-grade" OR "tenth grade" OR "eleventh-grade" OR "eleventh grade" OR "twelfth-grade" OR "twelfth grade" OR "1st grade" OR "2nd grade" OR "3rd grade" OR "4th grade" OR "5th grade" OR "6th grade" OR "7th grade" OR "8th grade" OR "9th grade" OR "10th grade" OR "11th grade" OR "12th grade" OR Mainstream OR "general education" OR "regular education")	
S10 S8 OR S9	848, 153
S11 **Design Terms** **Field Code: All text** "single-subject" OR SCD OR SCRD OR "single-case" OR Abab OR reversal OR "reversal design*" OR withdrawal OR "withdrawal design*" OR Baseline N20 treatment OR Baseline N20 intervention OR Baseline N20 game* OR Baseline N20 implement* OR "Alternating treatment*" OR multi*element OR multielement OR "multiple baseline" OR "multiple baseline design" OR "changing criterion" OR "changing criterion design" OR "group contingenc*"	11, 938
Final Search S1 AND S4 AND S10 AND S11	**105**

3) PsycInfo Search

EBSCO host

Table A- 3

Search Strategy Applied and Records Identified Across the PsycInfo Database

Search	Number of Records
S1 **Game Terms** **Field Code:** All Text (TX) TX (Game* OR gamif* OR "class-wide function-related intervention teams" OR cw-fit) OR DE Games	48,017
S2 **Behaviour Terms** **Field Codes:** Title, Abstract, Keywords, Subjects TI AB KW SU Behavio#r OR disrupt* OR conduct OR Management N5 classroom OR discipline OR attendance OR truancy OR absenteeism OR absence OR lateness OR punctuality OR timeliness OR tardiness OR engagement OR preparedness OR participation OR "active student responding" OR "on-task" OR "off-task" OR "hand-rais*" OR handrais* OR Tootl* OR tattl* OR homework OR "talking-out" OR "talking without permission" OR "seat-leaving" OR "out-of-seat" OR Aggression	1,518,463
S3 **Behaviour terms** **Exact Subject Terms** (identified by typing "classroom behavio#r" into thesaurus and using all relevant terms that appeared in list- recommended by librarian) DE "Classroom Behavior Modification" OR DE "Classroom Behavior" OR DE "Classroom Discipline" OR DE "Classroom Management" OR DE "Student Engagement"	16,749

Search	Number of Records
S4 S2 OR S3	1,518,463
S5 **School/Grade Terms** **Subject terms 1:** typed 'kindergarten' into thesaurus and selected relevant terms returned DE "Kindergarten Students" OR DE "Kindergartens"	7880
S6 **School/grade terms** **Subject terms 2:** typed 'elementary school" into thesaurus and selected relevant terms returned DE "Elementary School Students" OR DE "Primary School Students" OR DE "Intermediate School Students" OR DE "Elementary Schools" OR DE "Elementary Education" OR DE "School Based Intervention"	67755
S7 **School/grade terms** **Subject terms 3:** typed 'high school" into thesaurus and selected relevant terms returned DE "Junior High School Students" OR DE "High School Students" OR DE "High School Education" OR DE "Middle School Students" OR DE "Junior High Schools" OR DE "High Schools" OR DE "Secondary Education"	58,697
S8 S5 OR S6 OR S7	123,878
S9 **Field Codes:** Title, Abstract, Keywords, Subjects (want it to be broad but not too broad)	771,267

Search	Number of Records
Class* OR School* OR kindergarten OR "K-12" OR "Grade 1" OR "Grade 2" OR "Grade 3" OR "Grade 4" OR "Grade 5" OR "Grade 6" OR "Grade 7" OR "Grade 8" OR "Grade 9" OR "Grade 10" OR "Grade 11" OR "Grade 12" OR "first-grade" OR "second-grade" OR "Third-grade" OR "Fourth-grade" OR "fifth-grade" OR "sixth-grade" OR "seventh-grade" OR "eighth-grade" OR "ninth-grade" OR "tenth-grade" OR "eleventh-grade" OR "twelfth-grade" OR "1st grade" OR "2nd grade" OR "3rd grade" OR "4th grade" OR "5th grade" OR "6th grade" OR "7th grade" OR "8th grade" OR "9th grade" OR "10th grade" OR "11th grade" OR "12th grade" OR Mainstream OR "general education" OR "regular education"	
S10 S8 OR S9	772,273
S11 **Design Terms** **Field Code:** TX (all text) TX ("single-subject" OR SCD OR SCRD OR "single-case" OR Abab OR reversal OR "reversal design*" OR withdrawal OR "withdrawal design*" OR Baseline N20 treatment OR Baseline N20 intervention OR Baseline N20 game* OR Baseline N20 implement* OR "Alternating treatment*" OR multi*element OR multielement OR "multiple baseline" OR "multiple baseline design" OR "changing criterion" OR "changing criterion design" OR "group contingenc*")	95,458
Final Search S1 AND S4 AND S10 AND S11	**243**

4) Scopus Search

Table A-4

Search Strategy Applied and Records Identified Across the Scopus Database

Search	**Number of Records**
S1: TITLE-ABS-KEY (game* OR gamif* OR "class-wide function-related intervention teams" OR cw-fit)	306,340
S2: TITLE-ABS-KEY (behavior OR disrupt* OR conduct OR (management W/5 classroom) OR discipline OR attendance OR truancy OR absenteeism OR absence OR lateness OR punctuality OR timeliness OR tardiness OR engagement OR preparedness OR participation OR "active student responding" OR "on-task" OR "off-task" OR "hand-rais*" OR handrais* OR tootl* OR tattl* OR homework OR "talking-out" OR "talking without permission" OR "seat-leaving" OR "out-of-seat" OR aggression)	6, 743, 690
S3: TITLE-ABS-KEY (class* OR school* OR kindergarten OR "K-12" OR "Grade 1" OR "Grade 2" OR "Grade 3" OR "Grade 4" OR "Grade 5" OR "Grade 6" OR "Grade 7" OR "Grade 8" OR "Grade 9" OR "Grade 10" OR "Grade 11" OR "Grade 12" OR "first-grade" OR "second grade" OR "Third-grade" OR "Fourth-grade" OR "fifth-grade" OR "sixth-grade" OR "seventh-grade" OR "eighth-grade" OR "ninth-grade" OR "tenth-grade" OR "eleventh grade" OR "twelfth-grade" OR "1st grade" OR "2nd grade" OR "3rd grade" OR "4th grade" OR "5th grade" OR "6th grade" OR "7th grade" OR "8th grade" OR "9th grade"	5, 461, 985

Search	**Number of Records**
OR "10th grade" OR "11th grade" OR "12th grade" OR mainstream OR "general education" OR "regular education")	
S4: TITLE-ABS-KEY ("single-subject" OR scd OR scrd OR "single-case" OR abab OR reversal OR "reversal design*" OR withdrawal OR "withdrawal design*" OR (baseline W/20 treatment) OR (baseline W/20 intervention) OR (baseline W/20 game*) OR (baseline W/20 implement*) OR "Alternating treatment*" OR multi-element OR multielement OR "multiple baseline" OR "multiple baseline design" OR "changing criterion" OR "changing criterion design" OR "group contingenc*")	618, 906
S1 AND S2 AND S3 AND S4	309

5) Web of Science

Figure A-1

Screenshot Web of Science Search Page **(unable to download search)**

Set	Results	Save History / Create Alert · Open Saved History	Edit Sets	AND OR Combine	Select All ✕ Delete
# 5	263	#4 AND #3 AND #2 AND #1 *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years*	Edit		
# 4	373,068	TS=("single-subject" OR SCD OR SCRD OR "single-case" OR Abab OR reversal OR "reversal design*" OR withdrawal OR "withdrawal design*" OR Baseline NEAR/20 treatment OR Baseline NEAR/20 intervention OR Baseline NEAR/20 game* OR Baseline NEAR/20 implement* OR "Alternating treatment*" OR multi-element OR multielement OR "multiple baseline" OR "multiple baseline design" OR "changing criterion" OR "changing criterion design" OR "group contingenc*") *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years*	Edit		
# 3	3,619,575	TS=(Class* OR School* OR kindergarten OR "K-12" OR "Grade 1" OR "Grade 2" OR "Grade 3" OR "Grade 4" OR "Grade 5" OR "Grade 6" OR "Grade 7" OR "Grade 8" OR "Grade 9" OR "Grade 10" OR "Grade 11" OR "Grade 12" OR "first-grade" OR "second-grade" OR "Third-grade" OR "Fourth-grade" OR "fifth-grade" OR "sixth-grade" OR "seventh-grade" OR "eighth-grade" OR "ninth-grade" OR "tenth-grade" OR "eleventh-grade" OR "twelfth-grade" OR "1st grade" OR "2nd grade" OR "3rd grade" OR "4th grade" OR "5th grade" OR "6th grade" OR "7th grade" OR "8th grade" OR "9th grade" OR "10th grade" OR "11th grade" OR "12th grade" OR Mainstream OR "general education" OR "regular education") *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years*	Edit		
# 2	6,391,788	TS=(Behavior OR behaviour OR disrupt* OR conduct OR Management NEAR/5 classroom OR discipline OR attendance OR truancy OR absenteeism OR absence OR lateness OR punctuality OR timeliness OR tardiness OR engagement OR preparedness OR participation OR "active student responding" OR "on-task" OR "off-task" OR "hand-rais*" OR handrais* OR Tootl* OR tattl* OR homework OR "talking-out" OR "talking without permission" OR "seat-leaving" OR "out-of-seat" OR Aggression) *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years*	Edit		
# 1	242,258	TS=(Game* OR gamif* OR "class-wide function-related intervention teams" OR cw-fit) *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years*	Edit		

Appendix B: Microsoft Excel Worksheet used in Applying the WWC Design Standards

Note that this worksheet was developed by the researcher and is based on the WWC Study Review Guide (WWC, 2016) such that some items were copied directly to ensure compliance with the standards.

	Study Citation	Study ID	N Cases	N Participants	N Experiments	Describe	Intervention
	Design Quality						
Standards		Select from Dropdown				Information	
Preliminary	Does the study use a reversal/withdrawal, MBD, multiple probe, alternating treatment, changing criterion or other design? If yes, describe the study design and unit of analysis.	Yes/No					
Design Standard #1	Is the independent variable (i.e. the intervention) systematically manipulated with the researcher determining when and how the independent variable conditions change?	Yes/No					
Design Standard #2a	For all participants, did the authors collect IOA for each eligible outcome at least once in each phase?	Yes/No/Uncertain					
Design Standard #2b	Did the authors collect IOA on at least 20% of all sessions?	Yes/No/Uncertain					
Design Standard #2c	Did the authors collect IOA in at least 20% of the sessions in each condition? (i.e. baseline condition, intervention condition- author query not needed on this if unclear, but if if clear that this standard is NOT met, then the study does not meet standards)	Yes/No/Uncertain					
Design Standard #2d	Does IOA meet minimum acceptable values on each relevant outcome variable (i.e. 80% percentage agreement or .60 by Cohen's Kappa)? Overall agreement **for each outcome** much meet the minimum requirements. IOA needs to meet these minimum values for each outcome across all phases/cases but not seperately for each phase/case. Note how IOA is documented and the numerical values.	Yes/No/Uncertain					
Design Standard 2 FINAL	Is there at least one eligible outcome that meets WWC outcome requirements?	Yes/No/Uncertain					
Design Standard #3	Does the design include **at least three attempts to demonstrate an intervention effect at three different points in time?** (note, there may be separate cases, some of which meet this standard and some of which do not. This should be documented)	Yes/No/Uncertain					
Design Standard #4 (Reversal/Withdrawal)	Are there a sufficient number of data points in each phase? **Reversal/withdrawal design**s must have at least 5 data points per phase to Meet WWC Single-Case Design Standards without Reservations or 3 or 4 data points per phase to Meet WWC Single-Case Design Standards with Reservations.	Yes/No/Uncertain					
Design Standard #4 (MBD)	Are there a sufficient number of data points in each phase? **Multiple baseline and multiple probe designs** must have at least 5 data points per phase to Meet WWC Single-Case Design Standards without Reservations or 3 or 4 data points per phase to Meet WWC Single-Case Design Standards with Reservations.	Yes/No/Uncertain					
Design Standard #4 (ATD)	Are there a sufficient number of data points? **Alternating treatment designs** must have at least 5 data points per condition and at most 2 consecutive data points per condition to Meet WWC Single-Case Design Standards without Reservations. They must have at least 4 data points per condition and at most 2 consecutive data points per condition to Meet WWC Single-Case Design Standards with Reservations. Designs including multiple intervention comparisons (e.g., A vs. B, A vs. C, C vs. B) are rated separately for each comparison.	Yes/No/Uncertain					
Design Standard #4 (Changing Criterion)	Are there a sufficient number of data points in each phase? **Changing criterion designs** must have at least 5 data points per phase to Meet WWC Single-Case Design Standards without Reservations or 3 or 4 data points per phase to Meet WWC Single-Case Design Standards with Reservations.	Yes/No/Uncertain					
CONFOUND	Is the study free from any other confounds that may impact the design standard?	Yes/No/Uncertain					
	Design Quality Decisions by Experiment						
Experiment Number	Description (Participant, Outcome, Design)	Decision	Reason for Decision (Brief)	Standard not Met (If experiment does not meet standards)			
		Select Rating					
		Select Rating					
		Select Rating					
		Select Rating					
		Select Rating					
	Overall Study Rating						
Study Number	Rating (highest rating given to an experiment within the study)						
	Select Rating						

Adapted from: What Works Clearinghouse. (2016). *Study Review Guide Instructions for Reviewing Single-Case Designs Studies*. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_srg_scd_instructions_s3_v2.pdf

Items based on version 4 of the What Works Clearinghouse Design Standards: 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

Appendix C: Experiments included in effect size calculations by study

Table C-1

Experiments included in effect size calculations by study

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Casados (2012)	2	401		Classes 1-4 DC, Disruption, MBD	DB, DC classes	0.68	0.12	DB, DC classes	1.40	0.70
		402		Classes 1-4, PLC, Disruption, MBD	DB PLC classes	0.69	0.12	DB PLC classes	1.43	0.71
Conklin et al., (2017)	37 (13 experiments for which data could not be extracted)	601		2nd Grade, On-task Behaviour, ABABABABAB	OT, Compliance Combined	0.71	0.09	OT, Compliance Combined	3.41	0.70
		602		Kindergarten, On-Task Behaviour, ABABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Conklin et al., (2017) continued		603		7th Grade (am), On-task Behaviour, ABAB	OT, Compliance Combined continued			OT, Compliance Combined continued		
		604		7th Grade (pm), On-task Behaviour, ABAB						
		605		2nd grade, compliance, ABABABAB						
		606		Kindergarten, Compliance, ABAB						
		607		7th Grade (am), compliance, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		608		7th Grade (pm), compliance, ABAB	OT, Compliance Combined continued			OT, Compliance Combined continued		
		609		2nd grade P1 (triangle), On-task Behaviour, ABABABABAB	OT, individuals	0.65	0.07	OT, individuals	1.54	0.31
		610	The design of this experiment is ABABABABAB but only ABABAB was extracted as the data was difficult to	2nd Grade P2 (Closed circle), On-task Behaviour, ABABABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Conklin et al., (2017) continued			extract accurately		OT, individuals continued			OT, individuals continued		
		611		2nd Grade P3 (open circle), On-task Behaviour, ABABABABAB						
		614		Kindergarten P3 (open circle), On-task Behaviour, ABABAB						
		616		Kindergarten P5 (Open square), On-task Behaviour, ABABAB						
		618		7th Grade (am) P1 (closed circle),						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Conklin et al., (2017) continued				On-task Behaviour, ABAB	OT, individuals continued			OT, individuals continued		
		619		7th Grade (am) P2 (open circle), On-task Behaviour, ABAB						
		621		7th Grade (pm) P2 (open circle), On-task Behaviour, ABAB						
		622		2nd grade P1 (triangle), Hand raising, ABABABAB	Hand raising, individuals	0.63	0.08	Hand raising, individuals	1.31	0.28
		623	The design of the experiment is	2nd Grade P2 (closed circle),						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
			ABABABAB, but only ABABAB data extracted because validity was potentially compromised after that point (potentially not 3 data points in last A phase but hard to tell where data point should be at session 34)	Hand raising, ABABAB	Hand raising, individuals continued			Hand raising, individuals continued		
Conklin et al., (2017) continued		624		2nd Grade P3 (open circle), Hand Raising, ABABABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		628		Kindergarten P4, Hand raising, ABAB	Hand raising, individuals continued			Hand raising, individuals continued		
		629		Kindergarten P5 (open square), Hand raising, ABAB						
		631		7th Grade (am) P1 (closed circle), Hand raising, ABAB						
		632		7th Grade (am) P2 (open circle), Hand raising, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Conklin et al., (2017) continued		634		7th Grade (pm) P2 (open circle), Hand raising, ABAB	Hand raising, individuals continued			Hand raising, individuals continued		
		635	AB1 starts at data point 12 and the first AB iteration not considered (only 2 data points first phase)	2nd grade P1 (triangle), Out-of-Seat, ABABAB	OOS, individuals	0.63	0.09	OOS, individuals	.81	0.39
		637		2nd Grade P3 (open circle), Out-of-Seat, ABABABAB						

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Conklin et al., (2017) continued		641		Kindergarten P4, Out-of-Seat, ABAB	OOS, individuals continued			OOS, individuals continued		
		644		7th Grade (am) P1 (closed circle), Out-of-Seat, ABAB						
		647		7th Grade (pm) P2 (open circle), Out-of-Seat, ABAB						
		648		2nd grade P1 (triangle), Talking-Out, ABABABAB	Talking out, individuals	0.76	0.07	Talking out, individuals	1.60	0.34
		649	This experiment should have	2nd Grade P2 (closed circle),						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
			two additional AB phases changes but the B3 phase was to ambiguous to extract the data	Talking-Out, ABAB						
		650		2nd Grade P3 (open circle), Talking-Out, ABABABAB	Talking out, individuals continued			Talking out, individuals continued		
		653		Kindergarten P3 (open circle), Talking-out, ABAB						
		654		Kindergarten P4, Talking-Out, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		655		Kindergarten P5 (open square), Talking-Out, ABAB	Talking out, individuals continued			Talking out, individuals continued		
		657		7th Grade (am) P1 (closed circle), Talking-Out, ABAB						
		660		7th Grade (pm) P2 (open circle), Talking-Out, ABAB						

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Dadakhodjaeva et al., (2019)	4	901		Classrooms A, B, C, Disruptive Behaviour, MBD Across Classrooms	DB, whole classes	0.73	0.13	DB, whole classes	3.21	1.15
Dadakhodjaeva et al., (2019) continued		902		Classrooms A, B, C, Academically Engaged Behaviour, MBD Across Classrooms	AE, whole classes	0.7	0.14	AE, whole classes	2.55	0.50
		903		Cory, Michele, Conner, Disruptive Behaviour, MBD across Participants	DB individuals	0.53	0.16	DB individuals	1.52	0.29

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		904		Cory, Michele, Conner, Academically Engaged Behaviour, MBD across Participants	AE, individuals	-0.01	0.17	AE, individuals	1.24	0.34
Donaldson et al., (2017)	11	1202		Laurel, Disruptive Behaviour, ABAB	DB, individuals	0.5	0.06	DB, individuals	0.97	0.22
		1203		Zelda, Disruptive Behaviour, ABAB						
		1204		Carl, Disruptive Behaviour, ABAB						
		1205		Elton, Disruptive Behaviour, ABAB						
		1206		Kiley, Disruptive Behaviour, ABAB						

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Donaldson et al., (2017) continued		1207		Tamara, Disruptive Behaviour, ABAB	DB, individuals continued			DB, individuals continued		
		1208		Hope, Disruptive Behaviour, ABAB						
		1209		Masie, Disruptive Behaviour, ABAB						
		1210		Lamar, Disruptive Behaviour, ABAB						
		1211		Shaun, Disruptive Behaviour, ABAB						
		1212		Tanya, Disruptive Behaviour, ABAB						
Donaldson et al., (2018)	4	1301	The phases where there were different	Class 1, Disruptive	DB	0.73	0.09	DB	1.39	0.77

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Donaldson et al., (2018) continued			implementers were collapsed as one 'GBG' phase. A= Baseline, B= Experimenter, C= Teacher, D = Student, X= Choice	Behaviour, ABACADAX	DB continued			DB continued		
		1302	Data only extracted from first 4 phases as the following baseline phase had only 2 data points.	Class 2, Disruptive Behaviour, ABAD						
		1303	Data only extracted from first 6 phases as	Class 3, Disruptive						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
			the following baseline phase had only 1 data point.	Behaviour, ABACAD	DB continued			DB continued		
		1304	Final baseline phase ignored, only 2 data points	Class 4, Disruptive Behaviour, ABADAC						
Donaldson et al., (2015)	6	1401		Class 1 Activity 1, Disruptive Behaviour, ABABAB	DB	0.73	0.08	DB	1.89	0.55
Donaldson et al., (2015) continued		1402		Class 1 Activity 2, Disruptive Behaviour, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		1403		Class 2, Activity 1, Disruptive Behaviour, ABAB	DB continued			DB continued		
		1405		Class 3, Activity 1, Disruptive Behaviour ABAB						
		1407		Class 4, Activity 1, Disruptive Behaviour, ABAB						
		1409		Class 5, Activity 1, Disruptive Behaviour, ABAB						
Fallon et al., (2018)	2	1501	Note that spaces could not be left as too ambiguous.	x 3 classes, Academic Engagement, MBD	AE	0.61	0.15	AE	2.14	0.78

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		1502		x 3 classes, Disruptive Behaviour, MBD	DB	0.51	0.16	DB	0.67	0.89
Ford (2017)	8	1701		Classroom 1, AEB, ABAB	AE	0.75	0.09	AE	4.01	0.49
		1703		Classroom 2, AEB, ABAB						
		1705		Classroom 3, AEB, ABACAB						
		1707		Classroom 4, AEB, ABACAB						
		1702		Classroom 1, DB, ABAB	DB	0.74	0.09	DB	2.24	0.37
		1704		Classroom 2, DB, ABAB						

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
		1706		Classroom 3, DB, ABACAB	DB continued			DB continued		
		1708		Classroom 4, DB, ABACAB						
Hansen et al., (2017)	1	1901	Evaluated as a MBD therefore last two phases in 2nd and 3rd grade classes not considered.	2nd, 3rd & 4th grades, On-task Behaviour, MBD across classrooms	OT	0.58	0.13	OT	1.53	0.33
Hartman & Gresham (2016)	1	2003		Ms. King's Class, Disruptive Behaviour, ATD	n/a as ATD			n/a as ATD		
Hernan et al., (2019)	6	2101		Classroom A, Academic Engagement, ATD	n/a as ATD			n/a as ATD		

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		2102		Classroom A, Off-task Behaviour, ATD						
		2103		Classroom A, Mobile Device Use, ATD						
		2104		Classroom B, Academic Engagement, ATD						
		2105		Classroom B, Off-task Behaviour, ATD						
		2106		Classroom B, Mobile Device Use, ATD						

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Hine et al., (2015)	8	2201		Leo, Latency, ABAB	Latency to OT	0.74	0.11	Latency to OT	2.02	0.71
		2203		Samantha, Latency, ABAB						
		2205		Jonathan, Latency, ABAB						
		2207		Jamie, Latency, ABAB						
		2202		Leo, On-task behaviour, ABAB	OT	0.74	0.12	OT	2.55	0.54
		2204		Samantha, On-Task Behaviour, ABAB						
		2206		Jonathan, On-task Behaviour, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		2208		Jamie, On-task Behaviour, ABAB	OT continued			OT continued		
Hirsch et al., (2016)	1	2301		2nd grade class, PE class engagement, ABAB	PE AE	0.72	0.19	n/a		
Hoff & Ervin (2013) Hoff & Ervin (2013) continued	2	2401	Teacher-directed phase counted in intervention phase when calculating effect sizes.	x3 Classrooms, Disruptive Behaviour, Multiple Baseline Across Participants	DB, whole classes DB, whole classes continued	0.42	0.16	DB, whole classes DB, whole classes continued	0.53	0.35

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
		2402		Target Students, Disruptive Behaviour, Multiple Baseline Across Participants	DB, individuals	0.59	0.14	DB, individuals	0.99	0.44
Kelshaw-Levering et al., (2000)	1	2701	R-all and RR+ phases (i.e.,B and C phases) combined for effect size computation as one phase.	2nd grade class, Disruptive Behaviour, ABACBC	DB	0.64	0.2	n/a		

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Kleinman & Saigh (2011)	3	2801		9th grade class, Talk/verbal disruption, ABAB	DB (X3 types combined)	0.74	0.161	n/a		
		2802		9th Grade class, Aggression/physical disruption, ABAB						
		2803		9th grade class, Seat leaving, ABAB						
Lynne et al., (2017)	6	2901		Classroom A, AEB, ABAB	AE	0.69	0.13	AE	1.41	0.51
		2903		Classroom B, AEB, ABAB						
		2905		Classroom C, AEB, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		2902		Classroom A, DB, ABAB	DB	0.7	0.13	DB	1.98	0.52
		2904		Classroom B, DB, ABAB						
		2906		Classroom C, DB, ABAB						
Mitchell (2014)	15 (but only 6 used in effect size calculations - Only the composite DB variable was included in the meta-	3301		Classroom A, Academic Engagement, ABAB	AE	0.56	0.11	AE	0.89	0.74
		3303		Classroom B, Academic Engagement, ABAB						
		3305		Classroom C, Academic						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Mitchell (2014) continued	analysis as this is a combination of the 3 types graphed elsewhere in the study. Therefore, 9 experiments which met the WWC standards are not included here.)			Engagement, ABAB	AE continued			AE continued		
		3302		Classroom A, Disruptive Behaviour, ABAB	DB	0.64	0.11	DB	1.43	1.01
		3304		Classroom B, Disruptive Behaviour, ABAB						
		3306		Classroom C, Disruptive Behaviour, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Mitchell et al., (2015)	2	3401		Classroom A, Disruptive Behaviour, ABAB	DB	0.69	0.14	n/a		
		3403		Classroom C, Disruptive Behaviour, ABAB						
Mitchem, Young & Benyo (2001)	4	3501		Whole Classes, On task, MBD	Graphed data not extracted			Graphed data not extracted		
		3502		Rich, Howei, John, On task, MBD						
		3503		Jay, Arvelo, Recerdo, On task, MBD						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		3504		Cody, Helena, Rebecca, On task, MBD						
Naylor et al., (2018)	1	3707		Whole Class, On-task, ABAB	OT	0.56	0.19	n/a		
Nolan et al., (2014)	3	3801	Could be considered an MBD but we decided to evaluate as 3 separate ABAB designs	Classroom 1, Problem Behaviour, ABAB	DB	0.72	0.12	DB	1.77	1.38
		3802		Classroom 2, Problem Behaviour, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Nolan et al., (2014) continued		3803		Classroom 3, Problem Behaviour, ABAB	DB continued			DB continued		
Patrick et al., (1998)	2	3901		Grades 5,4,6, Appropriate Acts, MBD across participants	Graphed data not extracted			Graphed data not extracted		
		3902		Grades 5,4,6, Inappropriate Acts, MBD across participants						
Radley et al., (2016)	9	4001	Could be considered an MBD but we decided to evaluate as 3	Class A, AEB, ABAB	AE	0.73	0.12	AE	3.85	0.47

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
			separate ABAB designs		AE continued			AE continued		
		4003		Class B, AEB. ABAB						
		4005		Class C, AEB, ABAB						
Radley et al., (2016) continued		4002		Class A, DB, ABAB	DB, Decibel (Combined)	0.69	0.1216	DB, Decibel (Combined)	1.71	0.64
		4004		Class B, DB, ABAB						
		4006		Class C, DB, ABAB						
		4007		Class A, Decibels, ABAB						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		4008		Class B, Decibels, ABAB	DB, Decibel (combined; continued)			DB, Decibel (combined; continued)		
		4009		Class C, Decibels, ABAB						
Robichaux & Gresham (2014)	1	4101	There is a comparison of student selected and mystery rewards incorporated here but we just combined as one and evaluated as a MBD.	Three Classes, Disruptive Behaviour, MBD	DB	0.65	0.14	DB	2.06	0.55

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
Speight (2018)	1	4301	Training and treatment phases considered as one phase.	Three Classes, On task behaviour, MBD	OT	0.65	0.12	OT	3.10	0.64
Tanol et al., (2010)	6	4401	Tau effect size calculated initially but omitted from meta-analysis based on the similarity between this and the individual student data.	Class One, Rule violations, ABACBC	Rule violations, group	0.74	0.14	n/a		
		4402		Class Two, Rule violations, ABACBC						

Author (Year)	No. Exp MS/MSR	Exp ID	Exp./Study Notes	Exp. Description	Tau Label	Tau	SE Tau	BC-SMD Label	BC-SMD	SE BC-SMD
		4403		John Rule violations, ABACBC	Rule violation, individuals	0.7	0.1	Rule violation, individuals	3.12	0.35
		4405		Walter, rule violations, ABACBC						
		4406		Russell, Rule violations, ABACBC						
		4407		Viktor, Rule violations, ABACBC						
Vidoni et al., (2014)	1	4601		Three Classes, Steps per min, MBD across classes	Steps per min	0.68	0.15	Steps per min	2.09	0.74

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Wills (2002)	4	4801		Bryce, Michelle, Austin, On Task, MBD	OT	0.6	0.13	OT, individuals	2.27	0.83
		4803		Vance, ABAB, On Task				n/a		
		4804		Bryce, Michelle, Austin, Inappropriate behaviour, MBD	DB	0.52	0.14	DB, individuals	1.22	0.50
		4806		Vance, ABAB, Inappropriate behaviour				n/a		

Author (Year)	**No. Exp MS/MSR**	**Exp ID**	**Exp./Study Notes**	**Exp. Description**	**Tau Label**	**Tau**	**SE Tau**	**BC-SMD Label**	**BC-SMD**	**SE BC-SMD**
Wright & McCurdy (2012)	4	4901		Classroom 1 (kindergarten), On-task Behaviour, ABAC	OT	0.68	0.15	n/a		
		4903		Classroom 2 (4th grade), On-task behaviour, ACAB						
Wright & McCurdy (2012) continued		4902		Classroom 1 (kindergarten), DB, ABAC	DB	0.69	0.14			
		4904		Classroom 2 (4th grade), DB, ACAB						

Note. Exp = Experiment, No. Exp MS/MSR= Number of experiments within the study meeting the WWC standards with or without reservations, Tau/BCSMD Label= Label given to the effect sizes on the forest plots in Chapter 2 listed after the Author and year to specify further detail about the effect size.

Appendix D: Recruitment Materials for Chapter 3, Study 1

Appendix D1: Plain Language Statement for Board of Management, Chapter 3 Study 1

DUBLIN CITY UNIVERSITY

Plain Language Statement for Board of Management

Research Title: The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Problematic Classroom Behaviours

Student Researcher: Clare Bohan, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, H245D, School of Nursing and Human Sciences, DCU

Email: Sinead.smyth@dcu.ie

Ph: 01700 7422

What is the research about and why is it being conducted?

The school is being asked to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study is being funded by a Career Enhancement Grant awarded to Dr Smyth through DCU. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a Master of Research programme in Dublin City University.

The purpose of this study is to apply a new classroom management intervention to a school population in order to reduce class disruption and increase school attendance and other proactive school-based behaviours such as homework completion. The participating school will be asked to allow the student researcher to apply this classroom management intervention to classroom situations with consent and permission from relevant teachers and the students themselves. The intervention is based on a Fantasy Football league model, in which students choose 2-4 classmates to 'play on their team' for the term. Students each serve as a team captain and may choose any combination of 2-4 classmates to form their team. Any one student could be included on all teams or just on their own team. This is consistent with the idea of Fantasy Football in that Fantasy Football managers have a whole pool of players available to them weekly and all managers have access to the same players.

Teams gain and lose points based on their classroom behaviours. Each student is the captain of their own team and have the chance to earn double points for their team as an incentive to perform in a desirable manner.

The results of the study will provide an insight into how a group league applied to the classroom may benefit students and teachers alike. It is hoped that students may become more motivated to perform and attend school and that subsequently teachers will be supported and foster a successful learning environment. Information on the outcome of the project will be made available to class teachers and it can then be shared with parents and students who may be interested.

What do I have to do if I take part?

Should the Board consent to allowing the researcher access to the school, the school principal will be asked to assist in choosing a class group suitable for participation. The researcher will be present during class periods observing specific behaviours which will be predetermined with the help the class teacher. The teacher will receive information on how to conduct the league and along with the researcher, will have the power to award and deduct points. Aspects of the class may be recorded by audio or visual recording in order to obtain reliable data. The teacher may also be asked to keep track of some behaviours, such as pupils' school attendance or homework completion.

Participating students will be asked to take part in all activities pertaining to the league and also to fill out a short questionnaire on completion of the league, in order to assess how satisfactory they found it in terms of effectiveness and enjoyment.

Do I have to take part?

Teachers and students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary and this school was selected through collaboration with XXXX (redacted due to potentially identifying information).

Will I benefit from taking part and are there any risks involved?

It is hoped that the league will have a positive effect on teachers and students taking part in that disruptive behaviours will be decreased and prosocial classroom behaviours will be increased. Should the intervention prove successful, all teachers in the school will be given an opportunity to learn how best to apply it to the classroom.

Risk involved with the project is based mainly around the age profile of the child participants. The student researcher and all others involved in the project such as any research assistants are fully Garda vetted through the university and will have made themselves familiar with the DCU child protection policies. They will also make themselves familiar with any child protection policies relevant to the particular school. Teachers will also be asked to fill out a number of questionnaires with regards classroom management and stress levels related to teaching. Should they experience any discomfort in answering these questionnaires, they will be encouraged to seek support from relevant services within the school.

Contact details for relevant support parties will be made available to all parties involved in the project, including the researcher's' own details, supervisor details and details of an independent party from the DCU research ethics committee.

How will data be handled?

Data will be collected by the researcher on various behaviours however names will not be assigned to the data and the class will be analysed as a whole. Data on each team will be available only in the classroom setting and should examples of the league points be used in the write up, names will be changed. Data will not be shared with any outside bodies subject to legal limitations. Teacher questionnaires will not require that the teacher include their name. Due to the small number of teachers being recruited (3-5) and the fact that demographic information on these teachers will be included in the write up, it is a possibility that the teachers may be identifiable by those who know them particularly in reports circulated in the school. Nonetheless, the name of the school will not be disclosed in the write up and at no point will any clear identifying information be shared in the write up by the student researcher. Data including teacher questionnaires and consent forms with teacher and student names will be retained by the researcher's supervisor, Dr Smyth for a period of up to five years.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on page 1.

If participants have concerns about this study and wish to contact an independent person, please contact: The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix D2: Board of Management Consent Form, Chapter 3 Study 1

DUBLIN CITY UNIVERSITY

Informed Consent Form- Board of Management

Research Title: The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Problematic Classroom Behaviours

Student Researcher: Clare Bohan, MRes Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

sinead.smyth@dcu.ie

The purpose of this research is to investigate whether the Student Team Achievement Reward (STAR) League has a positive effect on classroom behaviours such as disruption and attendance. The population of interest is a group of secondary school students and their teachers.

I am interested in having classes taking part in the above research project, conducted by Clare Bohan, post graduate student in the School of Nursing and Human Sciences, DCU and supervised by Dr Sinéad Smyth, School of Nursing and Human Sciences, DCU. The study is funded by a Career Enhancement Grant awarded to Dr Smyth. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement (or had it read to me)	*Yes/No*
I understand all of the information provided	*Yes/No*
I have had an opportunity to ask questions and discuss this study	*Yes/No*
I have received satisfactory answers to all of my questions	*Yes/No*
I understand that participation is completely voluntary and that I can withdraw myself and my class from the study at any point without penalty	*Yes/No*
I understand that students will take part as team captains and will	

choose other part taking students to compete for their team throughout the league *Yes/No*

I understand that teachers will be asked to take part also by allowing implementation of the league in their classroom setting *Yes/No*

I understand that the data collected during this study will be kept in a confidential location subject to legal limitations *Yes/No*

I understand that the information I provide will be used for the researcher's completion of a post graduate programme and may be submitted for publication *Yes/No*

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to take part in this research project.

Participant's Signature: _________________

Name in Block Capitals:_____________________________________

Witness: ____________________________________

Date: _________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Chief Investigator, Dr. Sinead Smyth, School of Nursing and Human Sciences, Dublin City University, Dublin 9. Ph: 01 700 5000 EXT: 7422

If participants have concerns about this study and wish to contact an independent person,

please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Appendix D3: Parental Information Sheet and PLS, Chapter 3 Study 1

Information Sheet- STAR League

What is this about?

On the next page you will find more information about the STAR League

It is a game that is being introduced in your child's class

It is meant to improve school attendance and behaviour in class.

Who is doing it?

Researchers at Dublin City University (DCU)

- Clare Bohan is completing a Masters in DCU, supervised by Dr Sinéad Smyth.
- They are interested in applying the STAR League to a classroom .

What is involved?

- Researchers from DCU will be in your child's class to watch and make note of behaviour (e.g. attendance, homework, detentions) .
- Pupils can win prizes for behaving well in class.

Does My Son/Daughter Have to Take Part?

- You child will only take part if you and they want to.
- You asked to sign a consent form if you want your child to take part

For more detailed information and researcher contact details, please see Plain Language Statement overleaf

DUBLIN CITY UNIVERSITY

Plain Language Statement for Parents/Guardians

Research Title: The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Problematic Classroom Behaviours

Student Researcher: Clare Bohan, MRes Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, H245D, School of Nursing and Human Sciences, DCU

Email: Sinead.smyth@dcu.ie

Ph: Ext 7422

What is the research about and why is it being conducted?

You are being asked to allow your son/daughter to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study is being funded by a Career Enhancement Grant awarded to Dr Smyth through DCU. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a Master of Research/PhD programme in Dublin City University.

The purpose of this study is to apply a new classroom management intervention to a school population in order to reduce class disruption and increase school attendance and other proactive school-based behaviours such as homework completion.

What is involved?

a) For your son/daughter

Your son/daughter will be asked to take part in a classroom league, the Student Team Achievement Rewards (STAR) league, in which they will have the opportunity to gain points towards prizes. The league is based on a Fantasy Football model in that students will each serve as a team captains and choose their own 2-4 student team members for a 7-13 week period. Each captain may choose any student combination in their class to be on their team, in the same way that Fantasy Football managers have access to the full pool of soccer players each week. It is therefore possible for one student to be a part of a number of teams. During this period, all points gained and lost by team members are accumulated to form a team total. Captains can earn double points for their own team as an incentive to perform in a desirable manner. No time outside of regular school hours will be involved for students taking part in this project and it will not interfere with the layout of their regular school day.

They may be asked for their mobile phone number so that we can send them brief text reminders that the league is in place. They may also be asked to create an email address on which they may be contacted or link to a smartphone app for the duration of the study. This will simply be used to remind your child about different elements of the league and to let them know that points are available for certain class periods. We may also provide them with updates on their progress in this way. Any email addresses or phone numbers provided will be stored in a password protected location (see data section on next page) and on creating an email address, you will be informed so that you can monitor it accordingly.

b) For you as a parent

As a parent you can also engage with the project. If you consent, you can provide us with a mobile number so that we can remind you and your child that the league is in place during a certain upcoming class period. Don't worry, your number will be saved on a password protected and encrypted computer and will only be used for the purposes of this study. Only the primary researcher (Clare) will have access to this number and it will only be used to send text messages and not to make calls or contact you in any other manner.

Does my son/daughter have to take part?

Students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary.

You are being asked for your permission for your son/daughter to take part in the research project and we also enclose a form asking for 'assent' on you son/daughter's part. This means that while you as a parent/guardian grant the permission, it is also important that the student participating expresses their approval at taking part.

Behaviour Recording

Trained researchers will be present during some class periods to record aspects of the students' behaviour. This simply means that the researcher will observe the class and take note of incidences of behaviour such as disruption (incidences of students being out of their seat without permission, or talking out without permission) and active engagement with the class in progress. Behaviour such as homework completion and school attendance will also be noted with assistance from the school records. Your son/daughter may be asked to keep track of some aspects of their behaviour themselves. For example, we may ask them to track their engagement with the class by recording with a tick or x if they are doing the appropriate task at the appropriate time. Students' names will not be attributed to any data collected, be it collected themselves or observed by the researchers. Therefore, names will not be included in any papers written on the project. Non-participating students (those not giving consent) will not have their behaviour recorded.

Are there any benefits from taking part and are there any risks involved?

Students taking part in the intervention may experience direct and indirect benefits.

It is hoped that school attendance will increase and that overall disruptive classroom behaviours will be reduced. This will leave more time for teacher instruction and enhance the learning experiences for all.

Risk involved relates in the most part to the competition element of the league. If your son/daughter has participated in sports, classroom grading procedures and other classroom prize-based leagues, for example a star chart, they will have already experienced a similar situation to the STAR league.

Your son/daughter may withdraw participation at any time and points accumulated for their team up to that point will be kept for the teams, however no further points will be added or deducted for their behaviour.

Should you son/daughter feel uncomfortable in any way during the course of the project, they may contact the designated liaison person in the school.

How will data be handled?

Data will be collected by the researcher on disruptive behaviour, school attendance and other relevant classroom behaviours, however names will not be assigned to the data and the class will be analysed as a whole. Data on each team will be available only in the classroom setting and should examples of the league points be used in the write up for illustrative purposes, names will be changed. Mobile phone numbers of parents and/or email addresses of students will be stored on a password protected and encrypted excel file on the researchers' desktop computer. This information will only be used for league updates during the school term and not for any other purpose. Data will not be shared with any outside bodies subject to legal limitations. Data including consent forms with student names will be retained by the student researcher's supervisor, Dr Smyth, for a period of up to five years.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on page 1.

If participants have concerns about this study and wish to contact an independent person,

please contact: The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix D4: Parental Consent, Chapter 3 Study 1

DUBLIN CITY UNIVERSITY

Informed Consent Form- Parents/Guardians

Research Title: The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Problematic Classroom Behaviours

Student Researcher: Clare Bohan, MRes Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

sinead.smyth@dcu.ie

The purpose of this research is to investigate whether the Student Team Achievement Reward (STAR) League has a positive effect on classroom behaviours such as disruption and attendance. The population of interest is a group of secondary school students.

I give permission for my child, _____________________ to take part in the above research project, conducted by Clare Bohan, post graduate student in the School of Nursing and Human Sciences, DCU and supervised by Dr Sinéad Smyth, School of Nursing and Human Sciences, DCU. The study is funded by a Career Enhancement Grant awarded to Dr Smyth. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement (or had it read to me)	*Yes/No*
I understand all of the information provided	*Yes/No*
I have had an opportunity to ask questions and discuss this study	*Yes/No*
I have received satisfactory answers to all of my questions	*Yes/No*
I understand that my child, as a team captain, will be involved in choosing a team of 2-4 fellow students to take part in the league for its duration	*Yes/No*
I understand that my child may be asked to provide their mobile number and/or create an email account to receive updates on the league	*Yes/No*

I understand that participation is completely voluntary and that

I or my child can withdraw from the study at any point without penalty Yes/No

I understand that the data collected during this study will be kept in a

confidential location subject to legal limitations Yes/No

I understand that the information I provide will be used for the researcher's

completion of a post graduate programme and may be submitted for

publication Yes/No

I am willing to provide my phone number in order to receive reminders

about the project and prompts about my child's school attendance Yes/No

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent for my child to take part in this research project.

Participant's Signature: ______________________

Parent/Guardian Signature:___________________

Parent/Guardian Phone Number (Provide if desired): _________________________

Name in Block Capitals:____

Date: _________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Principal Investigator/Academic Supervisor, Dr. Sinead Smyth, School of Nursing and Human Sciences, Dublin City University, Dublin 9. Ph: 01 700 5000 EXT: 7422

If participants have concerns about this study and wish to contact an independent person,
please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Appendix D5: Student Assent, Chapter 3 Study 1

DUBLIN CITY UNIVERSITY

Minor Assent Form

Research Title: The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Problematic Classroom Behaviours

Student Researcher: Clare Bohan, MRes Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

sinead.smyth@dcu.ie

We are doing a research study about improving classroom behaviours associated with secondary school classrooms. We will be asking you to participate in a classroom league in which you will select a group of 2-4 classmates to take part on your team. This league is based on the idea of fantasy football. Your team will gain and lose points for you throughout the term based on how you behave in class and you will receive prizes based on the amount of points you gather.

The league will take part over the whole term and the teams and individuals gathering the most points will get the opportunity to receive prizes. We hope that you will benefit in that you will gain more from school and if you feel you already perform well at school, then you will be rewarded for it.

Clare may help you to create an email address or online account where you can get updates about your points during the league. You will get help from Clare and your teachers while creating this. Clare may also ask for your mobile phone number so that she can send you updates about the league, points and prizes.

If you decide at any stage that you do not want to take part, you can inform Clare, the researcher, and she will let you know that you can withdraw and recommend some supports if you feel you need them. When the study is finished, Clare will be writing a report on what was learned from it and using this report to complete her university course. Your name or the name of your school will not be used in this report.

Your parents/guardians have read about the study and you can discuss with them whether you wish to take part or not and make a decision together.

If you decide that you want to take part, please sign your name below.

Participant Signature (your signature):________________________

Participant Name in Block Capitals (your name in capital letters):

____

Date: __________________________________

Appendix D6: Teacher PLS, Chapter 3 Study 1

DUBLIN CITY UNIVERSITY

Plain Language Statement for Teachers

Research Title: The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Problematic Classroom Behaviours

Study Title: The Caught Being Good Game: Is public or private posting more effective in reducing disruption and improving upon engagement in a secondary school classroom?

Student Researcher: Clare Bohan, MRes Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, H245D, School of Nursing and Human Sciences, DCU

Email: Sinead.smyth@dcu.ie

Ph: 01 700 7422

What is the research about and why is it being conducted?

You are being asked to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study is being funded by a Career Enhancement Grant awarded to Dr Smyth through DCU. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a Master of Research/PhD programme in Dublin City University.

The purpose of this study is to apply a new classroom management intervention to a school population in order to reduce class disruption and increase class engagement. Participating teachers will be asked to allow the student researcher to apply this classroom management intervention to their classroom and partake in training where they will learn to apply the intervention themselves. The intervention is a game entitled the 'Caught Being Good Game'. It involves dividing the class group into teams of two or more. The teacher wears a vibrating device which will prompt them to scan the classroom at different times throughout the class. If upon scanning the room each member of a team are 'on task' or following the class rules, that team receives a point. Teams with a certain number of points at the end of class are eligible for a prize.

The aim of the research is to compare two versions of the game; a version where points are recorded privately by the teacher and a version where points are recorded publicly on the board.

Full training in administering the game/intervention will be provided by the researcher and regular meetings will be organised to ensure any questions you may have on a week to week basis can be addressed.

The results of the study will provide an insight into how this game applied to the classroom may benefit students and teachers alike. It will also provide a valuable insight into different versions of the game. It is hoped that students may become more motivated to behave in class and that subsequently teachers will be supported and foster a successful learning environment. Information on the outcome of the project will be made available to class teachers and it can then be shared with parents and students who may be interested.

What do I have to do if I take part?

The researcher will be present during class periods observing specific behaviours which will be predetermined with the help of yourself, the class teacher. You will receive information on how to conduct the game and award points. Aspects of the class may be recorded by audio or visual recording in order to obtain reliable data. The teacher may also be asked to keep track of some behaviour. You will also be asked to complete a short questionnaire on completion of the project, letting us know your feedback on the intervention.

Do I have to take part?

Teachers and students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary and this school was selected through collaboration with the XXX (redacted due to potentially identifying information).

Will I benefit from taking part and are there any risks involved?

Teachers may withdraw from use of the intervention at any time and the intervention will be removed from the class setting.

Should teachers feel any discomfort or experience adversity during the course of the project, they may contact a trusted party in the school, such as the school principal.

How will data be handled?

Data will be collected by the researcher on disruptive behaviour and engagement however names will not be assigned to the data and the class will be analysed as a whole. Data on each team will be available only in the classroom setting and should examples of the game points be used in the write up, names will be changed. Data will not be shared with any outside bodies subject to legal limitations. Teacher questionnaires will not require that the teacher include their name. Due to the small number of teachers being recruited (3-5) and the fact that demographic information on these teachers will be included in the write up, it is a possibility that the teachers may be identifiable by those who know them particularly in reports circulated in the school. Nonetheless, the name of the school will not be disclosed in the write up and at no point will any clear identifying information be shared in the write up by the student researcher. This write-up will include the student researcher's thesis, and may also include journal article publications and conference presentations. Data including teacher questionnaires and consent forms with teacher and student names will be retained by the researcher's supervisor, Dr Smyth for a period of up to five years.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on page 1. ***If participants have concerns about this study and wish to contact an independent person, please contact:*** The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix D7: Teacher Consent, Chapter 3 Study 1

DUBLIN CITY UNIVERSITY

Informed Consent Form- Teachers

Research Title: The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Problematic Classroom Behaviours

Student Researcher: Clare Bohan, MRes Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

sinead.smyth@dcu.ie

The purpose of this research is to investigate whether the Student Team Achievement Reward (STAR) League has a positive effect on classroom behaviours such as disruption and attendance. The population of interest is a group of secondary school students and their teachers.

I am interested in having my class taking part in the above research project, conducted by Clare Bohan, post graduate student in the School of Nursing and Human Sciences, DCU and supervised by Dr Sinéad Smyth, School of Nursing and Human Sciences, DCU. The study is funded by a Career Enhancement Grant awarded to Dr Smyth. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement (or had it read to me)	*Yes/No*
I understand all of the information provided	*Yes/No*
I have had an opportunity to ask questions and discuss this study	*Yes/No*
I have received satisfactory answers to all of my questions	*Yes/No*
I understand that participation is completely voluntary and that I can withdraw myself and my class from the study at any point without penalty	*Yes/No*
I understand that the data collected during this study will be kept in a confidential location subject to legal limitations	*Yes/No*

I understand that the information I provide will be used for the researcher's

completion of a post graduate programme and may be submitted for

publication *Yes/No*

I have read and understood the information in this form. My questions and concerns have been answered by the researchers. Therefore, I consent to take part in this research project.

Participant's Signature: ________________

Name in Block Capitals:____

Witness:___________________________________

Date: ________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Chief Investigator, Dr. Sinead Smyth, School of Nursing and Human Sciences, Dublin City University, Dublin 9. Ph: 01 700 5000 EXT: 7422

If participants have concerns about this study and wish to contact an independent person,
please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Appendix E: Team Daily/Weekly Leaderboard for Use During the CBGG, Chapter 3 Study 1

Team Name 1								
Team Name 2								
Team Name 3								

Points Needed For Prize: _________

Appendix F: Preference Assessment used in Chapter 3 Study 1

The following is a list of prizes you may get for behaving well in Maths class. Please put a number beside each prize with number '1' being your favourite (the prize you'd most like to win for behaviour in class) to number '6' being your least favourite

Lunch Tokens	
Shop Tokens	
Good Notes	
School Cinema Pass	
Reading/colouring/relaxing time at the end of class	
Stationary	

Do you have any other ideas for prizes for behaving well in Maths class?

____

____

____

____

Appendix G: Teacher Game Outline Used During Teacher Training, Chapter 3 Study 1

<u>Teacher Script</u>

Step 1: Announce the Game and Rules

Introduce the game to the students and announce the class rules. Direct the students to where they can see the class rules posted on the wall. Give examples of rule following and rule violating.

Example: "Today we are going to play a game. I am going to divide the class into teams and teams can earn points for following the class rules. The class rules are (read out class rules)".

Step 2: Explain the Game

Divide the class into the relevant teams (this only has to be done on day one, students will know their team after that) and allow 1 minute to choose a team name (again, only on day one, teams can keep their names after that point). Write team names on the board (if playing public version).

Tell students that you will randomly choose when to award points and that when you do, only teams on which every student is following the rules will get a point. If one team member is breaking the rules when the teachers chooses to scan the room, the whole team loses out on a point.

Step 3: Explain How to Win/ Reward Procedure

Explain to students that they can win the game by their team getting the most points. Explain that any team getting over the amount of points pulled from the draw at the end of class is automatically a winner, but the team with the most points get to vote on the prize awarded.

Show the students that you have a box with different prize names in it. Explain that each day you will choose two prizes from the box (you can do this at the beginning of class so that they know what they are working towards) and that the winning team will vote on which of these two prizes the teams over the criteria get.

Step 4: Start the Game

Announce that the game has started and start your timer. Carry on regular classroom instruction and when the timer vibrates, scan the room and award points a) privately in your own notebook/on a page or b) publically on the board under each team name (depending on the phase the research is in)

Step 5: End the game, add up points, allow winning team to vote

You will need to leave approx. 3-4 minutes at the end of class to add up scores and reward the winning team (s). Announce the end of the game, and add up the points a) on the board or b) in your own notes. Draw a criteria from the bowl so you know how many points are needed an announce this to the class. Announce which team has won in the case of both teams scoring over the criteria or announce which team got the most points in case of both teams getting over the criteria. Allow the winning team to vote on a prize (by show of hands) for themselves/all winning teams. Give out the prizes and dismiss the class.

Teacher Checklist (for teacher & researcher)

Step	**Tick When Complete**
Teacher announces that the game will be played	
Teacher divides the class into teams/reminds the classes of the team divisions	
Rules are posted on the wall	
Tell students how they can earn points/Review the rules	
Remind students how they can win & that criteria of points needed to win is decided at the end of the class	
Draw two prizes from the prize pool- announce what prizes are available to vote on today	
Start the timer & announce that the game has started	
Scan the room when the timer vibrates and privately/publicly record the points earned by each team	
Announce end of game ~ 3-5 minutes before end of class period	
Draw a criteria and announce the winner(s) at the end of the game	
Allow the team with the most points to vote on the prize for the day	
Give winning team(s) the prize or announce when they will have access to the prize	

Appendix H: Teacher CBGG Checklist, Chapter 3 Study 1

Teacher Checklist

Step	**Tick When Complete**
Teacher announces that the game will be played	
Teacher divides the class into teams/reminds the classes of the team divisions	
Rules are posted on the wall	
Tell students how they can earn points/Review the rules	
Remind students how many points they need to get the grand prize	
Start the timer & announce that the game has started	
Scan the room when the timer vibrates and privately/publicly record the points earned by each team	
Announce end of game ~ 3-5 minutes before end of class period	
Announce team points and write them onto the scoreboard	
Remind students how many points they need to get a prize and encourage them to keep working towards it OR during final day of game, announce the winners	

Appendix I: Observation sheet used to Collect Data in Chapter 3 Study 1

Observation Sheet

Class: ________ Date:________ Observer Initials: ___________ Phase: ___________

Codes: Engagement: Active Engagement= AET Passive Engagement= PET

Disruption: Off Task Motor= OFM Off Task Verbal= OFV

Interval	Disruption	Engagement	Interval	Disruption	Engagement
1.1			11.1		
1.2			11.2		
1.3			11.3		
1.4			11.4		
2.1			12.1		
2.2			12.2		
2.3			12.3		
2.4			12.4		
3.1			13.1		
3.2			13.2		
3.3			13.3		
3.4			13.4		
4.1			14.1		
4.2			14.2		
4.3			14.3		
4.4			14.4		
5.1			15.1		
5.2			15.2		
5.3			15.3		
5.4			15.4		
6.1			16.1		
6.2			16.2		
6.3			16.3		
6.4			16.4		
7.1			17.1		
7.2			17.2		
7.3			17.3		
7.4			17.4		
8.1			18.1		
8.2			18.2		
8.3			18.3		
8.4			18.4		
9.1			19.1		
9.2			19.2		
9.3			19.3		
9.4			19.4		
10.1			20.1		
10.2			20.2		
10.3			20.3		
10.4			20.4		

Appendix J: IRP-15 Completed by Ms. Allen in Chapter 3 Study 1

Caught Being Good Game Rating (Intervention Rating Profile –15/IRP-15)

You are being asked to complete this profile as a rating of the behaviour game as a whole which has been run in your classroom setting over the past number of weeks. Please circle the number that best describes your agreement or disagreement with each statement using the scale below.

1=strongly disagree	2=disagree	3=slightly disagree	4=slightly agree	5=agree	6=strongly agree

1.	This was an acceptable intervention for the problem behaviour(s).	1	2	3	4	5	6
2.	Most teachers would find this intervention appropriate for behaviour problems in addition to the ones described.	1	2	3	4	5	6
3.	This intervention proved effective in helping to change the problem behavior(s) of the classroom.	1	2	3	4	5	6
4.	I would suggest the use of this intervention to other teachers.	1	2	3	4	5	6
5.	The classroom behaviour problem was severe enough to warrant use of this intervention.	1	2	3	4	5	6
6.	Most teachers would find this intervention suitable for the problem behaviour(s) described.	1	2	3	4	5	6

7.	I would be willing to use this intervention in the classroom setting.	1	2	3	4	5	6
8.	This intervention would *not* result in negative side effects for the students.	1	2	3	4	5	6
9.	This intervention would be appropriate for a variety of students.	1	2	3	4	5	6
10.	This intervention is consistent with those I have used in classroom settings before.	1	2	3	4	5	6
11.	The intervention was a fair way to handle problem behaviour in the classroom.	1	2	3	4	5	6
12.	This intervention was reasonable for the problem behaviour(s) described.	1	2	3	4	5	6
13.	I liked the procedures used in this intervention.	1	2	3	4	5	6
14.	This intervention was a good way to handle the behaviour problem(s).	1	2	3	4	5	6
15.	Overall, this intervention was beneficial	1	2	3	4	5	6

Did you prefer the version of the game where feedback was delayed or where feedback was immediate? (circle your answer)

Delayed Feedback Immediate Feedback

Please indicate any ways in which you feel we could improve the behaviour game or indicate some suggestions for future implementations:

____
____
____
____
____
____
____

Reference: 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

With modifications adapted from: 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.

Appendix K: Modified CIRP Completed by Ms. Allen's Mathematics Class in Chapter 3 Study 1

Classroom Game Questionnaire

Read the questions about the team behaviour game you played in Maths class and circle 'yes' or 'no' as your answer;

1. Did you like the game used in your classroom?

   Yes No

2. Did you like participating in the game?

   Yes No

3. Do you think other students would like to use the game?

   Yes No

4. Did you like the rewards earned during the game?

   Yes No

5. Do you think the game has helped you do better in Maths class?

   Yes No

6. Do you think the game was fair?

   Yes No

7. Do you think the game caused any problems for you?

   Yes No

8. Do you think the game caused any problems for your classmates?

   Yes No

9. Did you prefer the game when the teacher recorded the points secretly (secret game) or the version where the points were recorded on the board (public game)? Circle your answer.

Secret Game Public Game

Do you have anything else to write about the classroom game?

____

____

____

____

____

____

____

____

____

____

____

____

_

Reference: Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in School Psychology* (pp. 251–288).

With modifications adapted from 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.

Appendix L: REC Approval Letter

Ollscoil Chathair Bhaile Átha Cliath
Dublin City University

DCU

Ms Clare Bohan
School of Nursing and Human Sciences

12 January 2017

REC Reference:	**DCUREC/2016/203**
Proposal Title:	**The STAR League: An Investigation into the Effects of a Novel Group Contingency League on Classroom Behaviours 1**
Applicant(s):	**Ms Clare Bohan & Dr Sinéad Smyth**

Dear Clare,

Further to a full committee review, the DCU Research Ethics Committee approves this research proposal.

Materials used to recruit participants should note that ethical approval for this project has been obtained from the Dublin City University Research Ethics Committee.

Should substantial modifications to the research protocol be required at a later stage, a further amendment submission should be made to the REC.

Yours sincerely,

Dónal O'Gorman

Dr Dónal O'Gorman
Chairperson
DCU Research Ethics Committee

DCU
Research & Innovation

Taighde & Nuálaíocht Tacaíocht
Ollscoil Chathair Bhaile Átha Cliath,
Baile Átha Cliath, Éire

Research & Innovation Support
Dublin City University,
Dublin 9, Ireland

T +353 1 700 8000
F +353 1 700 8002
E research@dcu.ie
www.dcu.ie

Appendix M: Recruitment Materials for Chapter 3 Study 2

Appendix M1: Teacher PLS, Chapter 3 Study 2

DUBLIN CITY UNIVERSITY

Plain Language Statement for Teachers

Research Title: The Caught Being Good Game: A comparison of immediate and delayed feedback during a gamified classroom management strategy

Student Researcher: Clare Bohan, PhD Candidate School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

Email: Sinead.smyth@dcu.ie

Ph: 01 700 7422

What is the research about and why is it being conducted?

You are being asked to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study is being funded by a Career Enhancement Grant awarded to Dr Smyth through DCU. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a PhD programme in Dublin City University.

The purpose of this study is to apply a classroom management intervention to a school population in order to reduce class disruption and increase class engagement. Participating teachers will be asked to allow the student researcher to apply this classroom management intervention to their classroom and partake in training where they will learn to apply the intervention themselves. The intervention is a game entitled the 'Caught Being Good Game'. It involves dividing the class group into two or more teams. The teacher wears a vibrating device (e.g. a smartphone held in a pocket) which will prompt them to scan the classroom at different times throughout the class. If upon scanning the room each member of a team are 'on task' or following the class rules, that team receives a point. Teams with a certain number of points at the end of class/end of the week are eligible for a prize.

The aim of the research is to compare two versions of the game; a version where points are recorded privately by the teacher and a version where points are recorded publically on the board.

Full training in administering the game/intervention will be provided by the researcher and regular meetings will be organised to ensure any questions you may have on a week to week basis can be addressed.

The results of the study will provide an insight into how this game applied to the classroom may benefit students and teachers alike. It will also provide a valuable insight into different versions of the game. It is hoped that students may become more motivated to behave in class and that subsequently teachers will be supported and foster a successful learning environment. Information on the outcome of the project will be made available to class teachers and it can then be shared with parents and students who may be interested.

What do I have to do if I take part?

The researcher will be present during class periods observing specific behaviours which will be predetermined with the help of yourself, the class teacher. You will receive information on how to conduct the game and award points. Aspects of the class may be recorded by audio or visual recording in order to obtain reliable data. The teacher may also be asked to keep track of some behaviour, including keeping note of integrity in implementing the intervention. This simply refers to amount of steps in implementing the game which you put in place. Clare will also monitor this in order to monitor the integrity of the project. Your feedback on the intervention is very important so you will also be asked to complete a short questionnaire on completion of the project to make your feedback known to the researchers.

Do I have to take part?

Teachers and students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary and this school was selected through collaboration with the XXX (redacted due to potentially identifying infofmation). You may withdraw your class at any time during the research process.

Will I benefit from taking part and are there any risks involved?

You may benefit in that it is expected that the intervention will reduce disruption and increase engagement across the recruited class group. This may mean that there is subsequently more time for class instruction related to the subject at hand. There are no major risks involved, however, should teachers feel any discomfort or experience adversity during the course of the project, they may contact a trusted party in the school, such as the school principal.

How will data be handled?

Data will be collected by the researcher on disruptive behaviour and engagement however names will not be assigned to the data and the class will be analysed as a whole. Data on each team will be available only in the classroom setting and should examples of the game points be used in the write up, names will be changed.

Data on student behaviour during class and teacher data including demographics and questionnaire results will be included in the student researcher’s thesis write-up and elements of this thesis may be submitted for publication in academic journals and/or presented at relevant conferences. No identifying information on teachers or the part-taking class will be included in articles submitted for publication or conference presentations. Teacher questionnaires will not require that the teacher include their name. Due to the small number of teachers being recruited (3-5) and the fact that demographic information on these teachers will be included in the write up, it is a possibility that the teachers may be identifiable by those who know them particularly in reports circulated in the school. Nonetheless, the name of the school will not be disclosed in the write up and at no point will any clear identifying information be shared in the write up by the student researcher.

Data will not be shared with any outside bodies subject to legal limitations.

Data including teacher questionnaires and consent forms with teacher and student names will be retained by the researcher's supervisor, Dr Smyth for a period of up to five years.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on page 1.

If participants have concerns about this study and wish to contact an independent person, please contact: The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix M2: Teacher Consent Form, Chapter 3 Study 2

DUBLIN CITY UNIVERSITY

Informed Consent Form- Teachers

Research Title: The Caught Being Good Game: A comparison of immediate and delayed feedback during a gamified classroom management strategy

Student Researcher: Clare Bohan, PhD Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

sinead.smyth@dcu.ie

The purpose of this research is to investigate whether the Caught being Good Game has a positive effect on classroom behaviours such as disruption and engagement. The population of interest is a group of secondary school students and their teachers.

I am interested in having my class taking part in the above research project, conducted by Clare Bohan, post graduate student in the School of Nursing and Human Sciences, DCU and supervised by Dr Sinéad Smyth, School of Nursing and Human Sciences, DCU. The study is funded by a Career Enhancement Grant awarded to Dr Smyth. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement (or had it read to me) *Yes/No*

I understand all of the information provided *Yes/No*

I have had an opportunity to ask questions and discuss this study

(please note that researcher contact details are provided above) *Yes/No*

I have received satisfactory answers to all of my questions *Yes/No*

I understand that participation is completely voluntary and that

I can withdraw myself and my class from the study at any point

without penalty *Yes/No*

I understand that the data collected during this study will be kept in a

confidential location subject to legal limitations *Yes/No*

I understand that the information I provide will be used for the researcher's

completion of a post graduate programme and may be submitted for

publication *Yes/No*

I have read and understood the information in this form. My questions and concerns have been answered by the researchers. Therefore, I consent to take part in this research project.

Participant's Signature: ____________________

Name in Block Capitals:____

Witness:____

Date: ____________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Chief Investigator, Dr. Sinead Smyth, School of Nursing and Human Sciences, Dublin City University, Dublin 9. Ph: 01 700 5000 EXT: 7422

If participants have concerns about this study and wish to contact an independent person,
please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Appendix M3: Teacher PLS for Teacher Behaviour, Chapter 3 Study 2

Study Title: Gamification in Classroom Management: Influences on Teacher Behaviour

Principal Investigator/Data Controller: Clare Bohan, BSc.

School of Nursing and Human Sciences, DCU

clare.bohan22@mail.dcu.ie

(01) 700 6866

Research Supervisor: Dr. Sinéad Smyth

School of Nursing and Human Sciences, DCU

Sinead.smyth@dcu.ie

(01) 700 7422

What is this research about and why is it being conducted?

You have been invited to take part in this research based on your willingness to take part in a classroom management intervention study. This statement informs you of an additional study which you may or may not be interested in taking part in.

The way in which teachers interact with students is important for student satisfaction at school and student-teacher relationships in general. This study aims to investigate whether the classroom management intervention in which you and your class are taking part influences teacher-student interactions, both positive and negative. Additionally, we wish to assess whether providing simple feedback to teachers on these interactions and teaching them to self-monitor their interactions throughout class may influence teacher response to student behaviour. This research is proposed for two reasons; first, to assess whether the classroom management intervention is influencing teacher-student interactions in a positive way and secondly, to assess whether the rates of positive interactions can be increased through feedback and/or self-monitoring. Feedback and self-monitoring are described in more detail in the following section.

This research is important as many schools country-wide are attempting to move towards a positive behavioural approach to managing behaviour in classroom settings. A positive behavioural approach simply means that there is a focus on encouragement of positive student behaviour rather than focusing on punishing negative or unwanted behaviour. The study will therefore be potentially useful for teachers who wish to adopt a positive behavioural approach as it should align with their professional goals and values. This study will assist in uncovering whether the classroom intervention is causing increases in positive teacher behaviour in line with a positive behavioural approach, and if not, whether giving teachers feedback on their behaviour can help with this.

What will I have to do if I consent to taking part?

If you decide to consent to taking part in this study, you will implement the classroom management strategy which you have previously consented to taking part in as you have been trained to do. As well as monitoring student behaviour, your interactions with students will be monitored also. These interactions may include the amount of praise statements used, amount of negative statements used, amount of punitive classroom management strategies used, amount of positive classroom management strategies used throughout a class period and amount of steps taking in correctly implementing the intervention. The behaviour to be monitored will be discussed with you before observations begin and the researchers will be open to your suggestions as to which areas of behaviour should be monitored.

When teacher-student interactions have been monitored for a number of days, the researcher will arrange to meet with you on a daily or weekly basis to provide you with feedback on your monitored behaviour. Meetings will take place with your schedule in mind and will take between 10 and 30 minutes. Should you find this feedback helpful, a self-monitoring component may be introduced whereby Clare will teach you to monitor your own rates of behaviour throughout classes. Self-monitoring is a useful strategy which is used by many individuals to track behaviour. It is hoped that self-monitoring of your own behaviour and comparing this monitoring with the observations made by the researcher will assist in developing a sense of awareness in your classroom management behaviour. Self-monitoring can be continued when the research project has ended at your own discretion.

Are there any risks involved in taking part in this research project?

It is not envisaged that you will experience serious discomfort or harm throughout the research process greater than that experienced in everyday life as a teacher. It is however important to note that if the researcher monitors your behaviour and it is not in line with your expectations when presented to you at a feedback meeting, you may feel uneased. An example of this may be if praise statements are monitored and you feel your praise statements are occurring often during class. The researcher may find that your praise statements are lower than you expected them to be. This may cause unease for you and a sense of dissonance. This said, we would hope that you could take feedback on board as a positive opportunity to alter your behaviour if it is not in line with your expectations. The researcher will continue to monitor your behaviour if you wish for them to do so and will continue to give you feedback to help you meet your expectations. You have the right to withdraw from the study at any stage during the process. If you choose to withdraw, your behaviour will not be monitored anymore and you will not be asked to take part in any more feedback meetings based on your behaviour.

There is a small risk that other teachers in your school may be aware that you are taking part in the study, thereby making you identifiable to them in any written reports. This being said, recruitment for this study is taking place in private and other staff members need not know that you are taking part in the feedback intervention if you do not wish to tell them so. Other teachers may be taking part themselves, which again may lead to some teacher knowing that you are taking part. It is important that you are aware of this before you consent to taking part.

Are there any benefits to taking part in this study?

Receiving feedback on monitored behaviour is a commonly used tactic in everyday life to improve upon behaviours. For example, thousands of people use Fitbits or similar fitness trackers to improve upon their step count. The premise here is similar and it is hoped that with regular feedback, your positive behaviours towards students can be maintained if occurring at a high rate already or improved upon if they are not occurring at a rate you are content with. Classrooms adopting more positive strategies have been found to have less office discipline referrals and suspensions, and although it is not expected that these benefits will occur on such a large scale here, it is hoped that it will help teachers move in a positive direction towards large scale changes.

Personal Data – GDPR Compliance

What will happen to the data which is collected on my behaviour?

Your personal data will be processed based on GDPR compliance.

Data will be collected by the primary researcher and other trained observers during class periods pre-determined by you. The data collected on you will include a consent form with your name signed and data on your behaviour during class. This data will be collected via paper and pen or on a portable laptop device/iPad. The data will be brought immediately and directly back to DCU

where it will be stored in a locked cabinet in the researcher's office protected by swipe card access. Electronic copies of data will be stored on a password protected laptop and/or computer and on an online storage system (Google Drive) which will be password protected. Individual data files will also be password protected to add an additional layer of protection. Documentation with your name (the signed consent form) will be stored separately to documentation on your behaviour to ensure you cannot be identified from behavioural data alone.
Data will be used for the completion of a PhD thesis in the area of classroom management and behaviour, therefore data will be published in this finished thesis. Names will not be assigned to data in these publications and/or presentations and you will be referred to be a pseudonym or simply by a letter (Teacher A). The name of your school will also not be published. The data may also be presented at relevant research conferences in the area of education and may be published in academic journals. Again, data will not be assigned your real name in these instances.
Confidentiality of information is protected within the limitations of the law in that any illegal behaviour observed or disclosed must be reported.
Data will be destroyed after a maximum period of 5 years by the principal researcher and/or supervisor. It is kept until the PhD thesis is completed. Hard copies of data will be shredded and disposed of in a confidential waste bin on DCU campus and soft copies will be completely deleted.
Your rights under the GDPR act are as follows;

1. The right to be informed; 2. The right of access; 3. The right to rectification; 4. The right to erasure; 5. The right to restrict processing; 6. The right to data portability; 7. The right to object; 8. Rights in relation to automated decision making and profiling.

Further information on these rights can be obtained on a website developed by the Data Protection Commission; gdprandyou.ie and the following URL will direct you to a pdf outlining your rights under GDPR: http://gdprandyou.ie/wp-content/uploads/2018/03/Rights-of-Individuals-under-the-General-Data-Protection-Regulation.pdf

Please note the contact details below for the DCU Data Protection Officer
Mr. Martin Ward (data.protection@dcu.ie Ph: (01) 7005118 /(01) 7008257

Voluntary Participation
Please note that your participation in this research project is completely voluntary and you can withdraw at any time.

Support services are available to teachers throughout their participation in this research project and in general at the employee assistance and wellbeing service, **Inspire,** at freephone number: 1800 411 057

If participants have concerns about this study and wish to contact an independent person, please contact:
The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000, e-mail **rec@dcu.ie**

Appendix M4: Teacher Consent form for Teacher Behaviour Monitoring, Chapter 3 Study 2

Study Title: Gamification in Classroom Management: Influences on Teacher Behaviour

Principal Investigator/Data Controller: Clare Bohan, BSc.

School of Nursing and Human Sciences, DCU

clare.bohan22@mail.dcu.ie

(01) 700 6866

Research Supervisor: Dr. Sinéad Smyth

School of Nursing and Human Sciences, DCU

Sinead.smyth@dcu.ie

(01) 700 7422

The current research project is being conducted to further the understanding of a) a teacher-student interactions while applying a classroom management strategy and b) the potential effects of feedback meetings and self-monitoring on these interactions. It is being carried out for completion of a PhD thesis by principal investigator, Clare Bohan.

Please complete the following (Circle Yes or No for each question)

I have read the Plain Language Statement (or had it read to me) *Yes/No*

I understand the information provided *Yes/No*

I have had an opportunity to ask questions and discuss this study *Yes/No*

I have received satisfactory answers to all my questions *Yes/No*

I am aware that my interactions with students while teaching will be observed and that I will receive feedback on my behaviour from the research team

Yes/No

I am aware that I may be asked to attend meetings with the primary researcher to discuss my interactions with students during class *Yes/No*

I am aware that I may be asked to self-monitor my behaviour during class *Yes/No*

I understand that I have the right to withdraw from participation in this study at any stage *Yes/No*

Data is collected by observation of classroom practices during school hours. Your data will be secured in a locked and secure cabinet on DCU campus and electronic copies of data will be secured on a password protected computer in an encrypted file. All data protection is subject to legal limitations. Data will be retained for a period of up to 5 years after which it will be disposed of as outlined in the plain language statement.

Signature: I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to take part in this research project

Participants Signature: ____________________

Name in Block Capitals: ____________________

Witness: ____________________

Date: ____________________

Appendix M5: Parental Information Sheet and PLS, Chapter 3 Study 2

Information Sheet- Caught Being Good Game

What is this about?

On the next page you will find more information about the Caught Being Good Game

It is a game that is being introduced in your child's class

It is meant to improve behaviour in class.

Who is doing it?

Researchers at Dublin City University (DCU)

- Clare Bohan is completing a PhD in DCU, supervised by Dr Sinéad Smyth.
- They are interested in applying the Caught Being Good Game to a classroom .

What is involved?

- Researchers from DCU will be in your child's class to watch and make note of behaviour (e.g. disruption, engagement)
- Pupils can win prizes for behaving well in class.

Does My Son/Daughter Have to Take Part?

- You child will only take part if you and they want to.
- You asked to sign a consent form if you want your child to take part

For more detailed information and researcher contact details, please see Plain Language Statement overleaf

DUBLIN CITY UNIVERSITY

Plain Language Statement for Parents/Guardians

Research Title: The Caught Being Good Game: A comparison of immediate and delayed feedback during a gamified classroom management strategy

Student Researcher: Clare Bohan, PhD Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

Email: Sinead.smyth@dcu.ie

Ph: Ext 7422

What is the research about and why is it being conducted?

You are being asked to allow your son/daughter to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study is being funded by a Career Enhancement Grant awarded to Dr Smyth through DCU. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a PhD programme in Dublin City University.

The purpose of this study is to apply a new classroom management intervention to a school population in order to reduce class disruption and increase other proactive school-based behaviours such as academic engagement.

What is involved?

Your son/daughter will be asked to take part in a classroom game, the Caught Being Good game, in which they will have the opportunity to gain points towards prizes. The game is team-based and students will work with their team to earn points. The class teacher will devise a set of class rules and teams will be asked to follow these rules as closely as they can. The teacher will scan the room at different times throughout the class and if a students team are all following the class rules, their team will receive a point. Teams reaching a certain number or criteria of point by the end of the last class of the week will be eligible for a prize. No time outside of regular school hours will be involved for students taking part in this project and it will not interfere with the layout of their regular school day.

Does my son/daughter have to take part?

Students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary.

You are being asked for your permission for your son/daughter to take part in the research project and we also enclose a form asking for 'assent' on you son/daughter's part. This means that while you as a parent/guardian grant the permission, it is also important that the student participating expresses their approval at taking part.

Behaviour Recording

Trained researchers will be present during some class periods to record aspects of the students' behaviour. This simply means that the researcher will observe the class and take note of incidences of behaviour such as disruption (incidences of students being out of their seat without permission, or talking out without permission) and active engagement with the class in progress. Behaviour such as homework completion and school attendance may also be noted with assistance from the school records. Your son/daughter may be asked to keep track of some aspects of their behaviour themselves. For example, we may ask them to track their engagement with the class by recording with a tick or x if they are doing the appropriate task at the appropriate time. Students' names will not be attributed to any data collected, be it collected themselves or observed by the researchers. Therefore, names will not be included in any papers written on the project. Non-participating students (those not giving consent) will not have their behaviour recorded.

Are there any benefits from taking part and are there any risks involved?

Students taking part in the intervention may experience direct and indirect benefits.

It is hoped that class engagement will increase and that overall disruptive classroom behaviours will be reduced. This will leave more time for teacher instruction and enhance the learning experiences for all.

Risk involved relates in the most part to the competition element of the league. If your son/daughter has participated in sports, classroom grading procedures and other classroom prize based leagues, for example a star chart, they will have already experienced a similar situation to the Caught Being Good Game.

Your son/daughter may withdraw participation at any time and their behaviour will not be recorded after that point. Data on their behaviour up to that point will be included in group analysis and will not be identifiable, therefore cannot be withdrawn.

Should you son/daughter feel uncomfortable in any way during the course of the project, they may contact the designated liaison person in the school.

How will data be handled?

Data will be collected by the researcher on disruptive behaviour, engagement and other relevant classroom behaviours, however names will not be assigned to the data and the class will be analysed as a whole. Data on each team will be available only in the classroom setting and should examples of the game points be used in the write up for illustrative purposes, names will be changed. Data on student behaviour during class will be included in the student researcher's thesis write-up and elements of this thesis may be submitted for publication in academic journals and/or presented at relevant conferences. No identifying information on the part-taking class will be included in articles submitted for publication or conference presentations. Data will not be shared with any outside bodies subject to legal limitations. Data including consent forms with student names will be retained by the student researcher's supervisor, Dr Smyth, for a period of up to five years.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on page 1.

If participants have concerns about this study and wish to contact an independent person,

please contact: The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix M6: Parental Consent form, Chapter 3 Study 2

DUBLIN CITY UNIVERSITY

nformed Consent Form- Parents/Guardians

Research Title: The Caught Being Good Game: A comparison of immediate and delayed feedback during a gamified classroom management strategy

Student Researcher: Clare Bohan, PhD Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

sinead.smyth@dcu.ie

The purpose of this research is to investigate whether the Caught Being Good Game has a positive effect on classroom behaviours such as disruption and engagement. The population of interest is a group of secondary school students.

I give permission for my child, ______________________ to take part in the above research project, conducted by Clare Bohan, post graduate student in the School of Nursing and Human Sciences, DCU and supervised by Dr Sinéad Smyth, School of Nursing and Human Sciences, DCU. The study is funded by a Career Enhancement Grant awarded to Dr Smyth. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement (or had it read to me) Yes/No

I understand all of the information provided Yes/No

I have had an opportunity to ask questions and discuss this study

(researchers details are included above for contacting with questions) Yes/No

I have received satisfactory answers to all of my questions Yes/No

I understand that my child will take part in a team-based game in their

classroom and will be eligible for prizes as part of that team Yes/No

I understand that participation is completely voluntary and that

I or my child can withdraw from the study at any point without penalty *Yes/No*

I understand that the data collected during this study will be kept in a

confidential location subject to legal limitations *Yes/No*

I understand that the information I provide will be used for the researcher's

completion of a post graduate programme and may be submitted for

publication *Yes/No*

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent for my child to take part in this research project.

Participant's (child's) Signature: ______________________

Parent/Guardian Signature:___________________

Parent Name in Block Capitals:____

Date: _________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Principal Investigator/Academic Supervisor, Dr. Sinead Smyth, School of Nursing and Human Sciences, Dublin City University, Dublin 9. Ph: 01 700 7422

If participants have concerns about this study and wish to contact an independent person,
please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Appendix M7: Student Assent form, Chapter 3 Study 2

DUBLIN CITY UNIVERSITY

Minor Assent Form

Research Title: The Caught Being Good Game: A comparison of immediate and delayed feedback during a gamified classroom management strategy

Student Researcher: Clare Bohan, PhD Candidate, School of Nursing and Human Sciences, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Nursing and Human Sciences, DCU

sinead.smyth@dcu.ie

We are doing a research study about improving classroom behaviours associated with secondary school classrooms. We will be asking you to participate in a classroom game in which you will gain points as part of a team in your class. Your team will gain and lose points for you throughout the term based on how you behave in class and you will receive prizes based on the amount of points you gather.

The game will take part during the whole term and during different weeks. Teams will get prizes if they gain enough points in one week of playing the game. We hope that you will benefit in that you will gain more from school and if you feel you already perform well at school, then you will be rewarded for it.

If you decide at any stage that you do not want to take part, you can inform Clare, the researcher, and she will let you know that you can withdraw and recommend some supports if you feel you need them. When the study is finished, Clare will be writing a report on what was learned from it and using this report to complete her university course. Your name or the name of your school will not be used in this report.

Your parents/guardians have read about the study and you can discuss with them whether you wish to take part or not and make a decision together.

If you decide that you want to take part, please sign your name below.

Participant Signature(your signature):________________________

Participant Name in Block Capitals (your name in capital letters):_______________________________

Date: ___________________

Appendix N: Preference Assessment used in Chapter 3 Study 2

Here is a list of some prizes you might receive for doing and behaving well in class.

Rate the prizes from 1-6: 1 is your favourite prize and 6 is your least favourite.

School Cinema Pass	
Stationary	
Sweets	
Homework Pass	
Free time	
Shop Vouchers for School Shop	

Are there any other prizes you can think of that you would like for doing well in class?

____

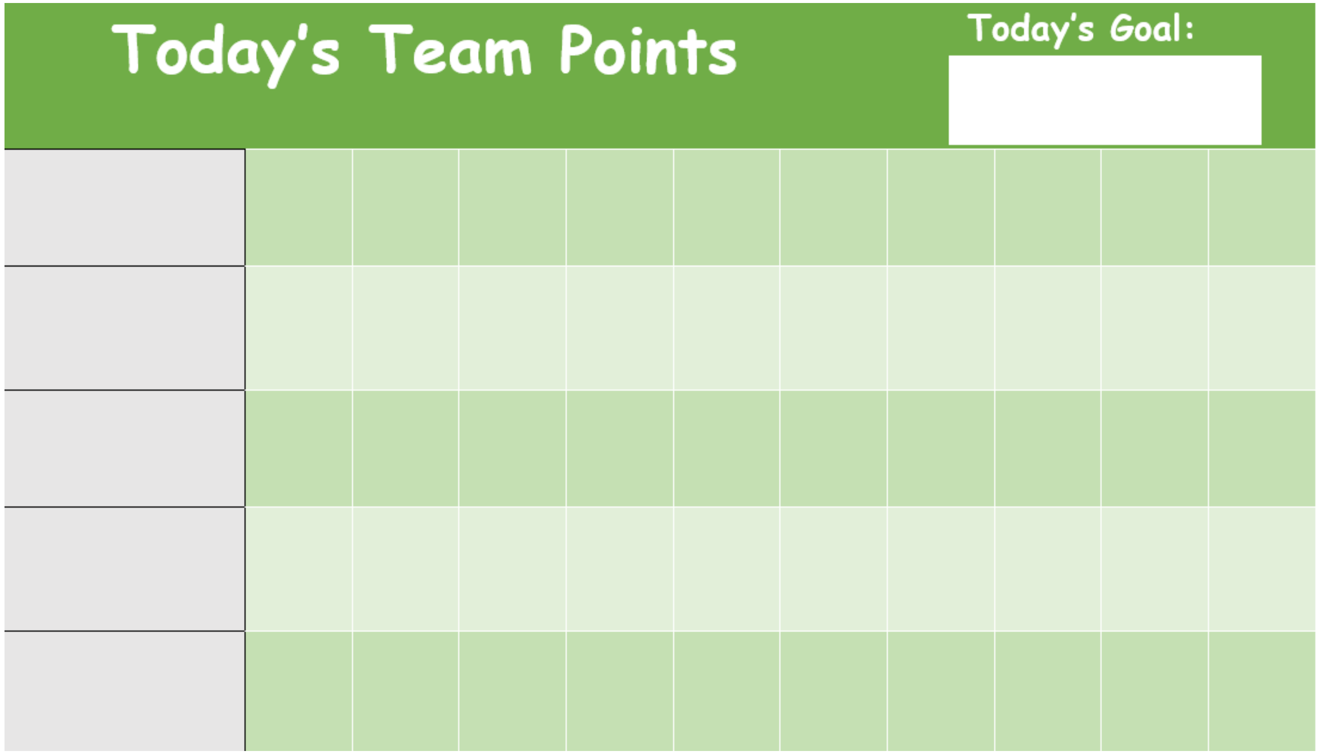
____

____

____

Appendix O: Scoreboards used in Study 2

Appendix O1: Daily scoreboard Study 2



Appendix O2: Weekly Scoreboard Study 2

This Week's Points

This Week's Goal: ______ This Week's Prize: ______

Team 1																														
Team 2																														
Team 3																														
Team 4																														
Team 5																														

Appendix P: Teacher CBGG Checklist Used in Chapter 3 Study 2

CBGG Checklist

Step	**Tick When Complete**
Teacher announces that the game will be played today	
Reminds the classes of the team divisions	
Rules are posted on the wall	
Review the rules	
Remind students that they earn points when their whole team is following rules when the teacher decides to scan the class	
Remind students how many points they need to get the grand prize (displayed on board)	
Start the timer & announce that the game has started	
Scan the room when the timer vibrates and privately/publically record the points earned by each team	
Announce end of game ~ 3-5 minutes before end of class period	
Announce team points and write them onto the weekly scoreboard	
Remind students how many points they need to get a prize and encourage them to keep working towards it OR during final day of game, announce the winners	

Appendix Q: Teacher Behaviour Data Collection Sheet Used in Chapter 3 Study 2

Observation Sheet

Class: _______Date:________ Observation Start time: __________ Finish Time:________ Observer Initials: _____

Frequency Count

Interval	Praise Statements	Negative Statements
1.1		
1.2		
2.1		
2.2		
3.1		
3.2		
4.1		
4.2		
5.1		
5.2		
6.1		
6.2		
7.1		
7.2		
8.1		
8.2		
9.1		
9.2		
10.1		
10.2		

Appendix R: BIRS Completed by Ms. Brady and Mr. Carroll in Chapter 3 Study 2

Caught Being Good Game Rating Scale for Teachers

Behaviour Intervention Rating Scale (BIRS; Elliot and Von Brock Treuting, 1991)

1=Strongly Disagree; ***2***=Disagree; ***3***=Slightly Disagree

4=Slightly Agree; ***5***=Agree; ***6***=Strongly Agree

1. This was an acceptable intervention for the students' problem behaviour(s).

1 2 3 4 5 6

2. Most teachers would find this intervention appropriate for behaviour problems in addition to the one(s) described (i.e. other behaviour issues in the classroom not targeted by the CBGG during this project).

1 2 3 4 5 6

3. The intervention proved effective in changing the students' problem behavior(s).

1 2 3 4 5 6

4. I would suggest the use of this intervention to other teachers.

1 2 3 4 5 6

5. The students' behaviour problem(s) were severe enough to warrant use of this intervention.

1 2 3 4 5 6

6. Most teachers would find this intervention suitable for the behaviour problem(s) described.

1 2 3 4 5 6

7. I would be willing to use this in the classroom setting again.

1 2 3 4 5 6

8. The intervention would ***not*** result in negative side-effects for students.

1 2 3 4 5 6

9. The intervention would be appropriate intervention for a variety of students.

1 2 3 4 5 6

10. The intervention is consistent with those I have used I have used in classroom settings.

1 2 3 4 5 6

11. The intervention was a fair way to handle the students' problem behaviour(s).

1 2 3 4 5 6

12. The intervention is reasonable for the behaviour problem(s) described.

1 2 3 4 5 6

13. I like the procedures used in the intervention.

1 2 3 4 5 6

14. The intervention was a good way to handle these students' behaviour problem(s).

1 2 3 4 5 6

15. Overall, the intervention was beneficial for the students.

1 2 3 4 5 6

16. The intervention quickly improved the students' behaviour.

1 2 3 4 5 6

17 .The intervention will produce a lasting improvement in the students' behaviour.

1 2 3 4 5 6

18. The intervention improved the students' behaviour to the point that it is not noticeably deviate from other students' behaviour (e.g. other well-behaved classes).

1 2 3 4 5 6

19. Soon after using the intervention, the teacher noticed a positive change in the problem behaviour.

1 2 3 4 5 6

20. The students' behaviour will remain at an improved level even after the intervention is discontinued.

1 2 3 4 5 6

21. Using the intervention should not only improve the students' behaviour in the classroom, but also in other settings (e.g., other classrooms, home).

1 2 3 4 5 6

22. When comparing these students with well-behaved peers before and after the use of the intervention, the students' and the peer's behaviours are more alike after using the intervention.

1 2 3 4 5 6

23. The intervention produced enough improvement in the students' behaviour so the behaviour no longer is a problem in the classroom.

1 2 3 4 5 6

24. Other behaviours related to the problem behaviour(s) were also improved by the intervention.

1 2 3 4 5 6

Did you prefer the CBGG **with delayed feedback** (recording the points in private and letting students know their score at end of class) or with **immediate feedback** (points recorded on the scoreboard throughout class)? **Circle your answer:**

Delayed Feedback Immediate Feedback

Why?

____

____

____

____

____

Do you have any further comments/feedback on the CBGG? If so, please record them below.

____

____

____

____

____

____

____

____

____

____

____

____

Adapted from: Elliott, S. N., & Von Brock Treuting, M. (1991). The Behavior Intervention Rating Scale: Development and Validation of a Pretreatment Acceptability and Effectiveness Measure. *The Journal of School Psychology, 29*, 43–51.

Adaptations include changing items to past tense and referring to a group of students rather than a 'child'.

Appendix S: DCU REC Approval for Teacher Behaviour Monitoring in Chapter 3 Study 2

Ollscoil Chathair Bhaile Átha Cliath
Dublin City University

Dr Sinéad Smyth
School of Nursing and Human Sciences

2nd November 2018

REC Reference:	**DCUREC/2018/178**
Proposal Title	**Gamification in Classroom Management: Influences on Teacher Behaviour**
Applicant(s):	**Dr Sinéad Smyth, Ms Clare Bohan & Dr Claire McDowell**

Dear Colleagues,

Further to expedited review, the DCU Research Ethics Committee approves this research proposal.

Materials used to recruit participants should note that ethical approval for this project has been obtained from the Dublin City University Research Ethics Committee.

Should substantial modifications to the research protocol be required at a later stage, a further amendment submission should be made to the REC.

Yours sincerely,

Dónal O' Gorman

Dr Dónal O'Gorman
Chairperson
DCU Research Ethics Committee

DCU
Research & Innovation

Taighde & Nuálaíocht Tacaíocht
Ollscoil Chathair Bhaile Átha Cliath,
Baile Átha Cliath, Éire

Research & Innovation Support
Dublin City University,
Dublin 9, Ireland

T +353 1 700 8000
F +353 1 700 8002
E research@dcu.ie
www.dcu.ie

Appendix T: Recruitment Materials Chapter 4

Appendix T1: Board of Management PLS, Chapter 4

DUBLIN CITY UNIVERSITY

Plain Language Statement for Board of Management Representative

Research Title: The Caught Being Good Game in a Primary School Classroom

Student Researcher: Clare Bohan, PhD Candidate, School of Psychology, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Psychology, DCU

Email: Sinead.smyth@dcu.ie

Ph: 01 700 7422

What is the research about and why is it being conducted?

The school is being asked to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a PhD programme in Dublin City University.

The purpose of this study is to apply a classroom management intervention to various class groups in order to reduce class disruption and increase engagement. The participating school will be asked to allow the student researcher to train a teacher or teachers in this classroom management intervention. The teacher(s) then apply it in the classroom. Consent and permission is obtained from relevant teachers, students parents, and the students themselves. The intervention, the Caught Being Good Game (CBGG) is a team-based game whereby students can earn points for their team by following class rules. The teacher divides the class into teams and students have the opportunity to name these teams as a group. Positively phrased class rules are devised by the teacher and posted in the classroom. When the game begins, the teacher sets a timer to vibrate at various intervals throughout the class. When signalled by the timer, the teacher scans the room and teams on which all students are following the rules receive a point. Teams reaching a certain criterion of points by the end of the game are eligible for prize.

The results of the study will provide an insight into how the CBGG applied to the classroom may benefit students and teachers alike. It is hoped that students may become more motivated to engage with class and that subsequently teachers will be supported and foster a successful learning environment. Information on the outcome of the project will be made available to class teachers and it can then be shared with parents and students who may be interested.

What do I have to do if I take part?

The school principal will be asked to assist in choosing a class group suitable for participation. The researcher will be present during class periods observing specific behaviours which will be predetermined with the help of the class teacher. The teacher will receive information on how to implement the CBGG and will have the responsibility of implementation and awarding of points throughout. Aspects of the class may be recorded by audio or visual recording in order to obtain reliable data. The teacher may also be asked to keep track of some behaviours, such as pupils' school attendance or homework completion.

Participating students will be asked to take part in all activities pertaining to the game and also to fill out a short questionnaire on completion of the game, in order to assess how satisfactory they found it in terms of effectiveness and enjoyment.

Do I have to take part?

Teachers and students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary and this school was identified for potential participation based on it's proximity to DCU and correspondence with the principal/relevant teachers.

Will I benefit from taking part and are there any risks involved?

It is hoped that the game will have a positive effect on teachers and students taking part in that disruptive behaviours will be decreased and prosocial classroom behaviours will be increased.

Risk involved with the project is based mainly around the age profile of the child participants. The student researcher and all others involved in the project such as any research assistants are fully Garda vetted through the university and will have made themselves familiar with the DCU child protection policies. They will also make themselves familiar with any child protection policies relevant to the particular school. Teachers will be fully supported in implementation of the intervention and will engage with regular meetings with the researcher where they can express and concerns and have any issues addressed.

Contact details for relevant support parties will be made available to all parties involved in the project, including the researcher's' own details, supervisor details and details of an independent party from the DCU research ethics committee.

How will data be handled?

Data will be collected by the researcher on various behaviours however names will not be assigned to the data. In the case where data is taken on an individual student, a pseudonym will be used when recording the data in the thesis or other written reports. Data on each team will be available only in the classroom setting and should examples of the game points be used in the write up, names will be changed. Data will not be shared with any outside bodies subject to legal limitations. Teacher questionnaires will not require that the teacher include their name. Due to the small number of teachers being recruited (2-5) and the fact that demographic information on these teachers will be included in the write up, it is a possibility that the teachers may be identifiable by those who know them particularly in reports circulated in the school. Nonetheless, the name of the school will not be disclosed in the write up and at no point will any clear identifying information be shared in the write up by the student researcher. Written reports on this research include the student's PhD thesis, journal publications and presentations at conferences. Data including teacher questionnaires and consent forms with teacher and student names will be retained by the researcher's supervisor, Dr Smyth for a period of up to five years.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on page 1. ***If participants have concerns about this study and wish to contact an independent person, please contact:*** The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix T2: Board of Management Consent form, Chapter 4

DUBLIN CITY UNIVERSITY

Informed Consent Form- Board of Management Representative

Research Title: The Caught Being Good Game in a Primary School Classroom

Student Researcher: Clare Bohan, PhD Candidate, School of Psychology, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Psychology, DCU

Email: Sinead.smyth@dcu.ie

Ph: 01 700 7422

The purpose of this research is to investigate whether the CBGG has a positive effect on classroom behaviours such as disruption and engagement. The population of interest is a group of primary school students and their teachers.

I am interested in having classes taking part in the above research project, conducted by Clare Bohan, post graduate student in the School of Psychology, DCU and supervised by Dr Sinéad Smyth, School of Psychology, DCU. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement (or had it read to me) *Yes/No*

I understand all of the information provided *Yes/No*

I have had an opportunity to ask questions and discuss this study (note, you may contact the student researcher at any time with questions and contact details are above) Yes/No

I have received satisfactory answers to all of my questions *Yes/No*

I understand that participation is completely voluntary and that

Any class may withdraw from the study at any point without penalty *Yes/No*

I understand that students will be asked to take part in the game, have

their behaviour monitored and be eligible to win prizes *Yes/No*

I understand that teachers will be asked to take part by allowing

implementation of the game in their classroom setting *Yes/No*

I understand that the data collected during this study will be kept in a

confidential location subject to legal limitations *Yes/No*

I understand that the information I provide will be used for the researcher's

completion of a post graduate programme and may be submitted for

publication *Yes/No*

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to take part in this research project.

Participant's Signature: ____

Name in Block Capitals:____________________________________

Witness: ____________________________________

Date: ________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Chief Investigator, Dr. Sinead Smyth, School of Nursing and Human Sciences, Dublin City University, Dublin 9. Ph: 01 700 7422

If participants have concerns about this study and wish to contact an independent person,

please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Appendix T3: Teacher PLS, Chapter 4

DUBLIN CITY UNIVERSITY

Plain Language Statement for Teachers

Research Title: The Caught Being Good Game in a Primary School Classroom

Student Researcher: Clare Bohan, PhD Candidate School of Psychology, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Psychology, DCU

Email: Sinead.smyth@dcu.ie

Ph: 01 700 7422

What is the research about and why is it being conducted?

You are being asked to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a PhD programme in Dublin City University.

The purpose of this study is to apply a classroom management intervention (The Caught Being Good Game; CBGG) to a school population in order to look at its effect on class disruption and class engagement. Participating teachers will be asked to partake in training on the CBGG and apply this in their classroom for a short period each day (e.g. 20 minutes to a half an hour). The CBGG involves dividing the class group into two or more teams. The teacher wears a vibrating device (e.g. a smartphone held in a pocket, or a smart wrist device like a Fitbit) which will prompt them to scan the classroom at different times throughout the class. If upon scanning the room each member of a team are 'on task' or following the class rules, that team receives a point. Teams with a certain number of points at the end of class and/or end of the week are eligible for a prize.

Full training in administering the CBGG will be provided by the researcher and regular meetings will be organised to ensure any questions you may have on a week to week basis can be addressed.

The results of the study will provide an insight into how this game applied to the primary level classroom may benefit students and teachers alike. It is hoped that students may become more motivated to behave in class and that subsequently teachers will be supported in fostering a successful learning environment.

What do I have to do if I take part?

The researcher will be present during class periods observing specific student and teacher behaviours. Specific class periods and behaviour which will be monitored will be predetermined with the help of yourself, the class teacher. You will receive training on how to conduct the game and award points. Aspects of the class may be recorded by audio or visual recording in order to obtain reliable data. The teacher may also be asked to keep track of some behaviour, including keeping note of integrity in implementing the intervention. This simply refers to amount of steps in implementing the game which you put in place. Clare will also monitor this in order to monitor the integrity of the project. Your feedback on the intervention is very important so you will also be asked to complete a short questionnaire on completion of the project to make your feedback known to the researchers.

Do I have to take part?

Teachers and students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary. You may withdraw your class at any time during the research process.

Will I benefit from taking part and are there any risks involved?

You may benefit in that it is expected that the intervention will reduce disruption and increase engagement across the recruited class group. This may mean that there is subsequently more time for class instruction. There are no major risks involved, however, should teachers feel any discomfort or experience adversity during the course of the project, they may contact a trusted party in the school, such as the school principal.

How will data be handled?

Data will be collected by the researcher on disruptive behaviour and engagement however names will not be assigned to the data. Data on each team will be available only in the classroom setting and should examples of the game points be used in the write up, names will be changed.

Data on student behaviour during class and teacher data including demographics and questionnaire results will be included in the student researcher's thesis write-up and elements of this thesis may be submitted for publication in academic journals and/or presented at relevant conferences. No identifying information on teachers or the part-taking class will be included in articles submitted for publication or conference presentations. Teacher questionnaires will not require that the teacher include their name. Due to the small number of teachers being recruited (2-5) and the fact that demographic information on these teachers will be included in the write up, it is a possibility that the teachers may be identifiable by those who know them particularly in reports circulated in the school. Nonetheless, the name of the school will not be disclosed in the write up and at no point will any clear identifying information be shared in the write up by the student researcher.

Data will not be shared with any outside bodies subject to legal limitations.

Data including teacher questionnaires and consent forms with teacher and student names will be retained by the researcher's supervisor, Dr Smyth for a period of up to five years.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on page 1.

If participants have concerns about this study and wish to contact an independent person, please contact: The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix T4: Teacher Consent Form, Chapter 4

DUBLIN CITY UNIVERSITY

Informed Consent Form- Teachers

Research Title: The Caught Being Good Game in a Primary School Classroom

Student Researcher: Clare Bohan, PhD Candidate, School of Psychology, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Psychology, DCU

sinead.smyth@dcu.ie

The purpose of this research is to investigate whether the Caught being Good Game has a positive effect on classroom behaviours such as disruption and engagement. The population of interest is a group of primary school students and their teachers.

I am interested in having my class taking part in the above research project, conducted by Clare Bohan, post graduate student in the School of Psychology, DCU and supervised by Dr Sinéad Smyth, School of Psychology, DCU. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement (or had it read to me) *Yes/No*

I understand all of the information provided *Yes/No*

I have had an opportunity to ask questions and discuss this study

(please note that researcher contact details are provided above) *Yes/No*

I have received satisfactory answers to all of my questions *Yes/No*

I understand that participation is completely voluntary and that

I can withdraw myself and my class from the study at any point

without penalty *Yes/No*

I understand that the data collected during this study will be kept in a

confidential location subject to legal limitations *Yes/No*

Continued overleaf, turn the page…

I understand that the information I provide will be used for the researcher's completion of a post graduate programme and may be submitted for publication *Yes/No*

I have read and understood the information in this form. My questions and concerns have been answered by the researchers. Therefore, I consent to take part in this research project.

Participant's Signature: _________________

Name in Block Capitals:____

Witness:_____________________________________

Date: _________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Chief Investigator, Dr. Sinead Smyth, School of Psychology, Dublin City University, Dublin 9. Ph: 01 700 7422

If participants have concerns about this study and wish to contact an independent person, please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Appendix T5: Parental PLS, Chapter 4

Information Sheet: The Caught Being Good Game

What is this about?

On the next page you will find more information about a project on the Caught Being Good Game.

It is a game that is being introduced in your child's class.

It is meant to improve behaviour in class.

Who is doing it?

Researchers at Dublin City University (DCU)

- Clare Bohan is completing a PhD in DCU, supervised by Dr Sinéad Smyth.
- They are interested in applying the Caught Being Good Game to a classroom.

What is involved?

- Researchers from DCU will be in your child's class to watch and make note of behaviour (e.g. disruption, engagement)
- Pupils can win prizes for behaving well in class.
- Any notes on your child's behaviour will be kept completely anonymous. The name of your child, their school, or their teacher will not be included in anything that is written about this project.

Does My Son/Daughter Have to Take Part?

- All students may play the game in class with their teacher, but only those who returned signed consent forms will have their behaviour recorded.
- You are asked to sign a consent form if you want your child to take part.

For more detailed information and researcher contact details, turn this page.

Plain Language Statement for Parents/Guardians

Research Title: The Caught Being Good Game in a Primary School Classroom

Student Researcher: Clare Bohan, PhD Candidate, School of Psychology, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Psychology, DCU

Email: Sinead.smyth@dcu.ie

Ph: 01 700 7422

What is the research about and why is it being conducted?

You are being asked to allow your son/daughter to take part in a research study to be conducted by Clare Bohan and supervised by Dr Sinéad Smyth, Dublin City University. The study has been approved by the DCU Research Ethics Committee. This study is being conducted for submission in partial completion of a PhD programme in Dublin City University.

The purpose of this study is to apply a new classroom management strategy to a primary school population in order to reduce class disruption and increase other proactive school-based behaviours such as academic engagement.

What is involved?

Your son/daughter will be asked to take part in a classroom game, the Caught Being Good game, in which they will have the opportunity to gain points towards prizes. The game is team-based and students will work with their team to earn points. The class teacher will devise a set of class rules and teams will be asked to follow these rules as closely as they can. The teacher will scan the room at different times throughout the class and if a student's team are all following the class rules, their team will receive a point. Teams reaching a certain number of points by the end of the game-period and/or the end of the school week will be eligible for a prize. No time outside of regular school hours will be involved for students taking part in this project and it will not interfere with the layout of their regular school day. Students will have the chance to provide us with written feedback when the project is over in the form of a survey. This will be completed during class time with the help of the class teacher and students will not have to write their name on this survey so it will remain anonymous and confidential.

Does my son/daughter have to take part?

Students in selected classes are eligible to take part. Participation by class teachers and students is completely voluntary. The teacher may decide that the whole class will take part in the game as part of their own classroom management protocol, but ***unless you have given permission, your son/daughters behaviour will not be monitored*** during the game.

You are being asked for your permission for your son/daughter to take part in the research project and we will also ask for 'assent' on your son/daughter's part. This means that while you as a parent/guardian grant the permission, it is also important that the student participating expresses their approval at taking part.

Behaviour Recording

Trained researchers from DCU will be present during some class periods to record aspects of the students' behaviour. This simply means that the researchers will observe the class and take note of incidences of behaviour such as disruption (incidences of students being out of their seat without permission, or talking out without permission) and active engagement with the class in progress. Behaviour such as homework completion and school attendance may also be noted with assistance from the school records. Your son/daughter may be asked to keep track of some aspects of their behaviour themselves. For example, we may ask them to track their engagement with the class by recording with a tick or x if they are doing the appropriate task at the appropriate time. Students' names will not be attributed to any data collected, be it collected themselves or observed by the researchers. Therefore, names will not be included in any papers written on the project. Non-participating students (those not giving consent) will not have their behaviour recorded but may still take part in the game at the teacher's discretion.

Data will be looked at by researchers at the group and individual level. That is, we will look at whether the game is improving the behaviour of the class as a whole, as well as for individual students within the class. This is important as sometimes a game may be benefitting the whole class, but some individual students are not benefitting as much. If we see this is happening, we may be able to make some changes to the game to ensure every student is getting a chance at benefitting.

Are there any benefits from taking part and are there any risks involved?

Students taking part in the game may experience direct and indirect benefits. It is hoped that class engagement will increase and that overall disruption will be reduced. This will leave more time for teacher instruction and enhance the learning experiences for all. Risk involved relates in the most part to the competition element of the game. If your son/daughter has participated in sports, classroom grading procedures and other classroom prize-based strategies, for example a star chart, they will have already experienced a similar situation to the Caught Being Good Game. Your son/daughter may withdraw participation at any time and their behaviour will not be recorded after that point. Data on their behaviour up to that point that has been included in group analysis (in other words the whole class average) and will not be identifiable, therefore cannot be withdrawn. Should your son/daughter feel uncomfortable in any way during the course of the project, they may contact the designated liaison person in the school.

How will data be handled?

Data will be collected by the researcher on disruptive behaviour, engagement and other relevant classroom behaviours, however student names will not be assigned to the data. In the case of individual data, students will be assigned a pseudonym (fake name) or a number when the data is presented so that they are not identifiable. Data on each team (points earned etc.) will be available only in the classroom setting and should examples of the game points be used in the write up for illustrative purposes, names will be changed. The data will be written up for inclusion in Clare's PhD thesis and again, no reference will be made to any particular students' real name, the name of their teacher or school. Data may also be written up for publication in academic peer-reviewed journals and/or presented at research conferences.

Data will not be shared with any outside bodies subject to legal limitations. Data including consent forms with student names will be retained by the student researcher's supervisor, Dr Smyth, for a period of up to five years. Data will be stored in a soft copy on an encrypted Google Drive system and hard copy data will be stored in a locked cabinet in the researcher's office on DCU campus.

If you should have any questions either before consenting, or during the completion of the project, please contact either the project supervisor or student researcher on the contact details provided on the previous page.

If participants have concerns about this study and wish to contact an independent person,

please contact: The Secretary, Dublin City University Research Ethics Committee, c/o Research and Innovation Support, Dublin City University, Dublin 9. Tel 01-7008000

Appendix T6: Parental Consent Form, Chapter 4

DUBLIN CITY UNIVERSITY

Informed Consent Form- Parents/Guardians

Research Title: The Caught Being Good Game in a Primary School Classroom

Student Researcher: Clare Bohan, PhD Candidate, School of Psychology, DCU

Clare.bohan22@mail.dcu.ie

Supervisor: Dr. Sinead Smyth, School of Psychology, DCU

sinead.smyth@dcu.ie

The purpose of this research is to investigate whether the Caught Being Good Game has a positive effect on classroom behaviours such as disruption and engagement. The population of interest is a group of primary school students.

I give permission for my child, _____________________ to take part in the above research project, conducted by Clare Bohan, post graduate student in the School of Psychology, DCU and supervised by Dr Sinéad Smyth. This research has been granted ethical approval by the DCU Research Ethics Committee.

Please read the following statements and circle your answer;

I have read the Plain Language Statement/Information Sheet (or had it read to me) *Yes/No*

I understand all of the information provided *Yes/No*

I have had an opportunity to ask questions and discuss this study

*(**NOTE:** researchers' details are included above for contacting with questions)* *Yes/No*

I have received satisfactory answers to all of my questions *Yes/No*

I understand that my child will take part in a team-based game in their

classroom and will be eligible for prizes as part of that team *Yes/No*

I understand that participation is completely voluntary and that

I/my child can withdraw from the study at any point without penalty *Yes/No*

I understand that the data collected during this study will be kept in a

confidential location subject to legal limitations *Yes/No*

I understand that the information I provide will be used for the researcher's

completion of a post graduate programme and may be submitted for

publication and/or presented at conferences *Yes/No*

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent for my child to take part in this research project.

Participant's (child's) Signature: ______________________

Parent/Guardian Signature:___________________

Parent Name in Block Capitals:____

Date: _________________

If participants have further questions about this study or their rights, or if they wish to lodge a complaint or concern, they may contact:

The Principal Investigator/Academic Supervisor, Dr. Sinead Smyth, School of Psychology, Dublin City University, Dublin 9. Ph: 01 700 7422

If participants have concerns about this study and wish to contact an independent person,
please contact:

The Secretary, Dublin City University Research Ethics Committee, c/o Office of the Vice-President for Research, Dublin City University, Dublin 9. Tel 01-7008000

Hi, I'm Clare! I'd like to get your help on a project that I am doing!

What is the project all about?	**This project is about helping you do your school-work by playing a game. You will work with a team of your class-mates to earn points for doing your best and following your teachers instructions. If you earn enough points with your team, you will get rewards for doing so well. Your teacher will tell you all about how you can earn points and you can ask him/her about that too.**
Your Behaviour	Clare will sit in your classroom and watch your class for a little while each day. She will be watching to see if the game is helping you do better during class. She will take down some notes while she is there. Sometimes, Clare might bring someone with her to help her do this.
What will happen to the notes about my behaviour?	Any information that Clare takes during your class will be used to help her write a report for her college course. Don't worry, she won't use your name, the name of your teacher, or the name of your school in this report. Your name will be disguised to make sure that no-body knows it was written about you. Any notes taken by Clare will be kept safe in a locked drawer in DCU or on a laptop with a strong password.

What if I don't want you to watch my behaviour?	If you don't want Clare to watch you while she is in your class, that is OK. You do not have to take part. Your teacher might decide that you should still work with your team to earn points, but Clare will not watch you when she comes to visit your class.
What if I have some questions?	You can ask Clare some questions at a time that your teacher allows, or you can ask your teacher questions too. Your parents have read all the about the project as well so you can talk to them about it. Clare's email address is below and you can contact her with the help of your parents.

Clare's Email Address: clare.bohan22@mail.dcu.ie

Consent Form for Students

I have read all about the study and had the chance to ask questions.

I want to take part in the project (tick the box if this is ok) ☐

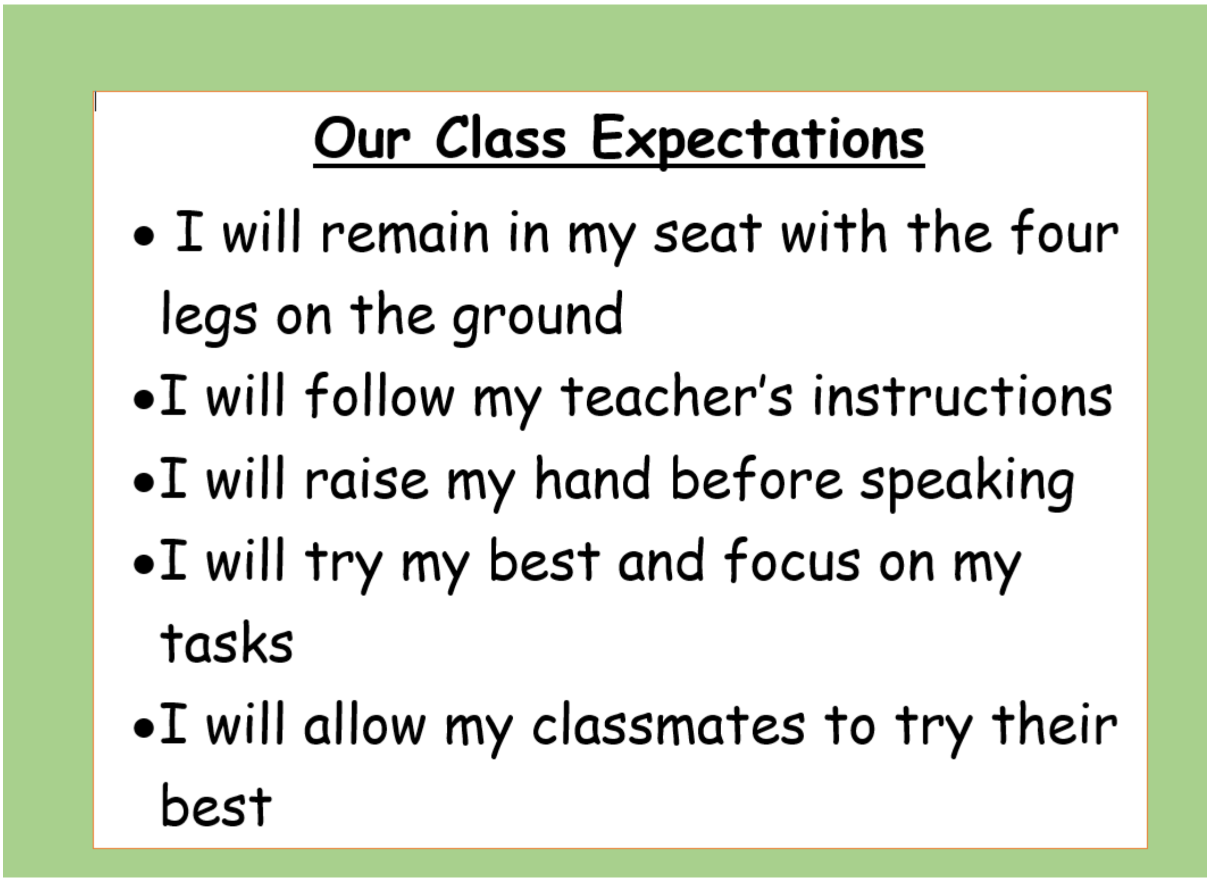
Sign your name here if you want to take part:___________________________

Write your name in capital letters here:

____

Date: ___________________________

Appendix U: Class Expectations for Fourth Class in Chapter 4



Appendix V: Score Board Used to Record Team Points in Chapter 4

Today's Team Points								Today's Goal:		

Appendix W: Teacher CBGG Checklist Used in Chapter 4

CBGG Checklist

Step	**Tick When Complete**
Teacher announces that the game will be played today	
Reminds the classes of the team divisions	
Rules are posted on the wall	
Review the rules	
Remind students that they earn points when their whole team is following rules when the teacher decides to scan the class (i.e. how to win the game)	
Remind students how many points they need to get the prize (displayed on board)	
Announce that the game has started	
Scan the room when the timer vibrates and record the points earned by each team	
Announce end of game after the last 'behaviour check' has been completed	
Total the team points for today's game	
Announce the winners and give access to the prize	

Appendix X: Data Collection Sheet for Student Behaviour, Chapter 4

20 minute observation

10 second intervals +

5 second record

Observation Sheet

Class: _______Date:________ Observation Start time: __________ Finish Time:________ Observer Initials: _____________

DB (partial interval recording) = Disruptive Behaviour, record EACH TYPE with + (occurrence) or – (non-occurrence)

AEB (momentary time sampling) = Academically engaged behaviour, record with + (occurrence) or – (non-occurrence)

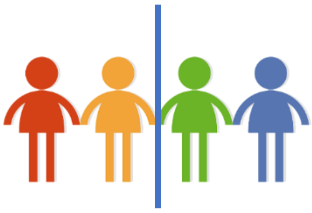
Interval	DB			AEB	Interval	DB			AEB
	TALK	OOS	MOT			TALK	OOS	MOT	
1.1					11.1				
1.2- TARGET 1					11.2- TARGET 1				
1.3					11.3				
1.4- TARGET 2					11.4- TARGET 2				
2.1					12.1				
2.2- TARGET 1					12.2- TARGET 1				
2.3					12.3				
2.4- TARGET 2					12.4- TARGET 2				
3.1					13.1				
3.2- TARGET 1					13.2- TARGET 1				
3.3					13.3				
3.4- TARGET 2					13.4- TARGET 2				
4.1					14.1				
4.2-TARGET 1					14.2- TARGET 1				
4.3					14.3				
4.4- TARGET 2					14.4- TARGET 2				
5.1					15.1				
5.2- TARGET 1					15.2- TARGET 1				
5.3					15.3				
5.4- TARGET 2					15.4- TARGET 2				
6.1					16.1				
6.2- TARGET 1					16.2- TARGET 1				
6.3					16.3				
6.4- TARGET 2					16.4- TARGET 2				
7.1					17.1				
7.2-TARGET 1					17.2- TARGET 1				
7.3					17.3				
7.4- TARGET 2					17.4- TARGET 2				
8.1					18.1				
8.2-TARGET 1					18.2- TARGET 1				
8.3					18.3				
8.4-TARGET 2					18.4- TARGET 2				
9.1					19.1				
9.2-TARGET 1					19.2- TARGET 1				
9.3					19.3				
9.4-TARGET 2					19.4- TARGET 2				
10.1					20.1				
10.2-TARGET 1					20.2- TARGET 1				
10.3					20.3				
10.4- TARGET 2					20.4- TARGET 2				

Hi, I'm Clare! I'd like to get your help on a project that I am doing!

What's involved with helping Clare?

To help with your school work,

your teacher will divide your class into teams to play a game.



Your team can earn points  for trying your best and following class rules.

Your team can earn prizes for earning lots of points.

Clare will sit in your classroom sometimes and watch your class. She will be checking to see if the game is helping you do better during class.

If you do not want Clare to watch your behaviour during class, you can put your hand up and tell the teacher. You may feel OK about Clare watching my behaviour another time, but it is OK if you do not want her to at all.

You can tick the box and write your name if you would like to take part.

I want to take part ☐

Signed (write your name):

____

Appendix Z: Class Expectations for Ms. Ellis' Class, Chapter 5

Our Class Expectations

- Look at and Listen to your teacher
- Hands up and wait for your teacher
- Do your best at your work
- Respect your friends & let them do their work
- Stay in your seat

Appendix AA: Scoreboard Used to Record Points in Chapter 5

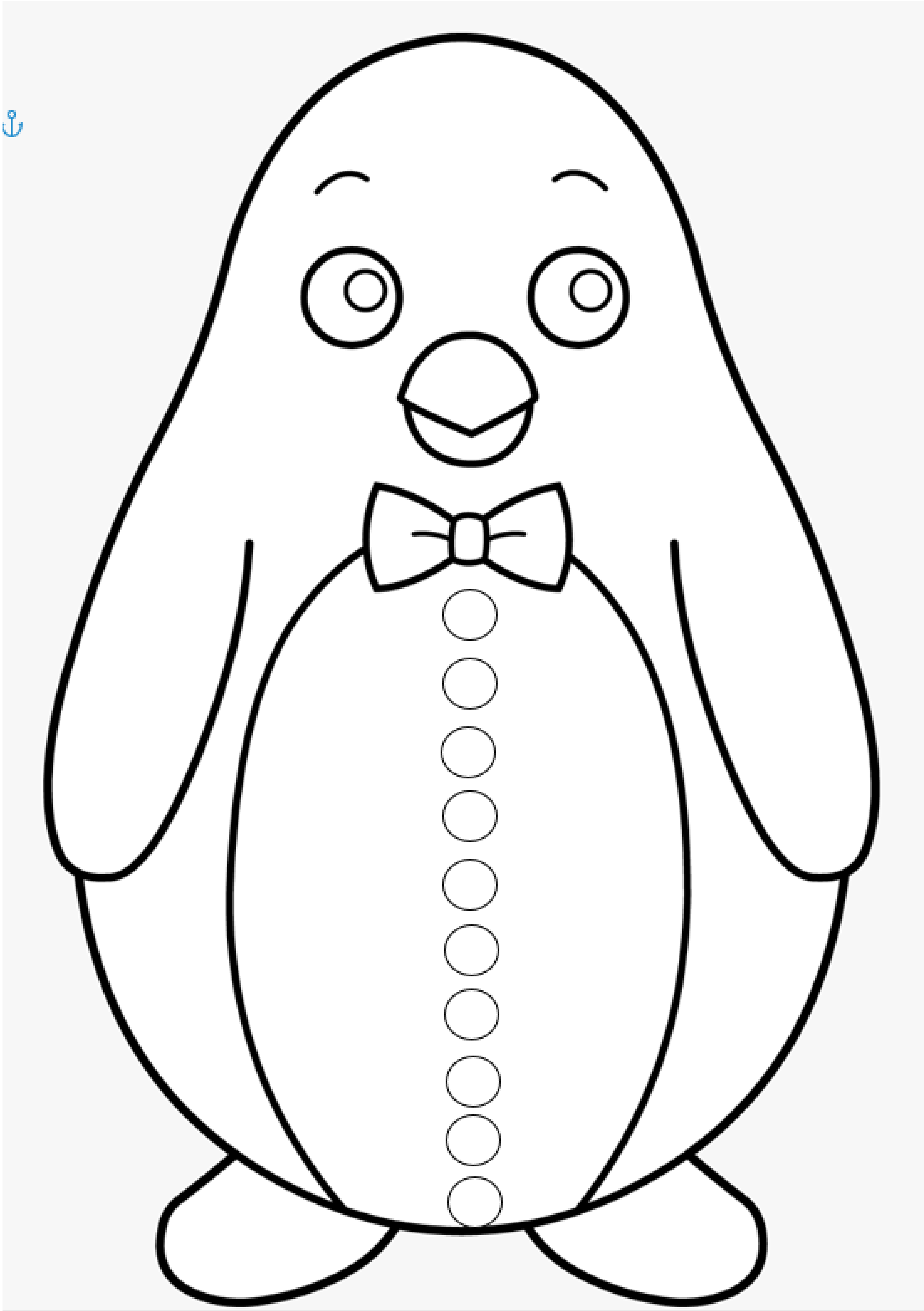
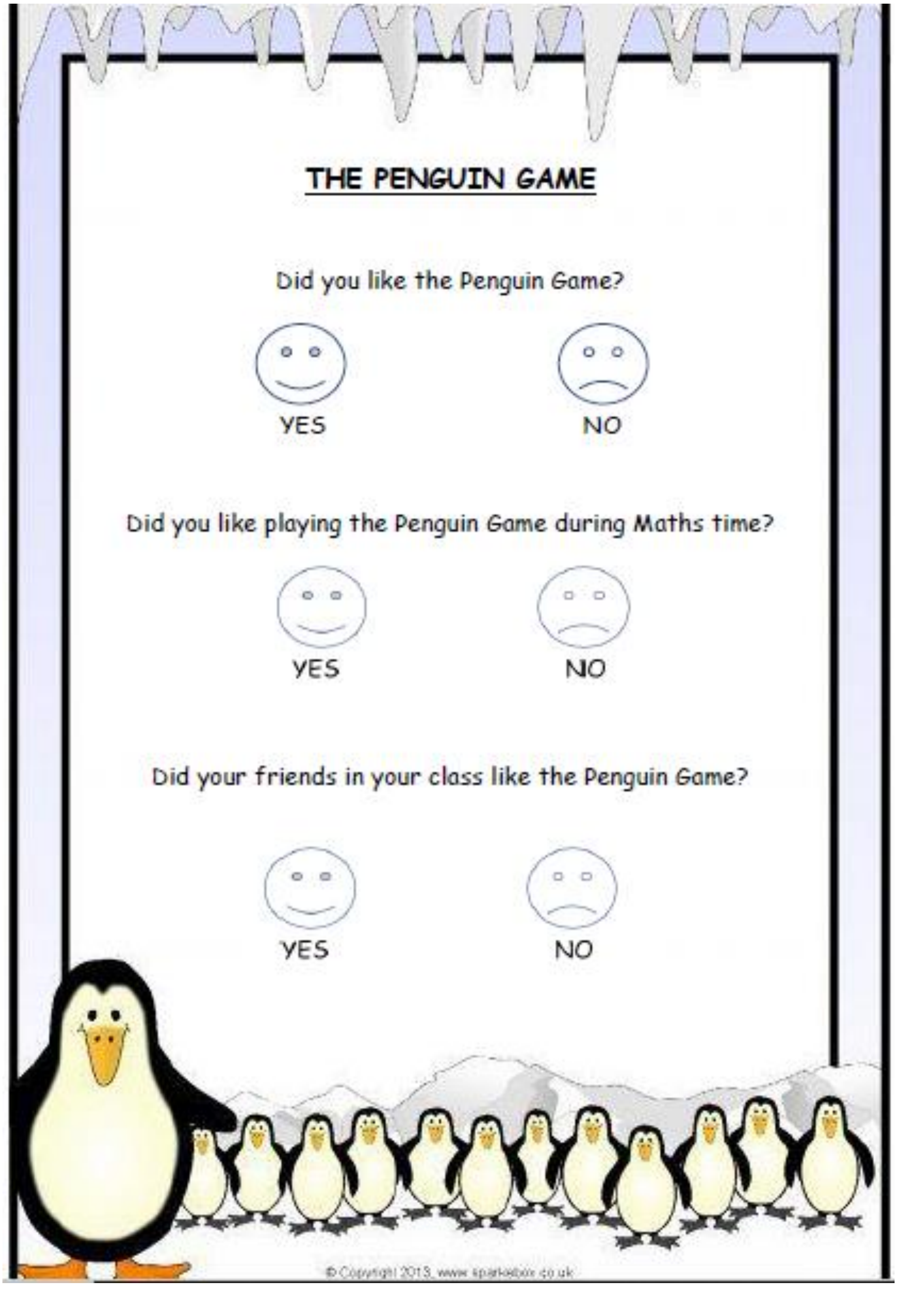
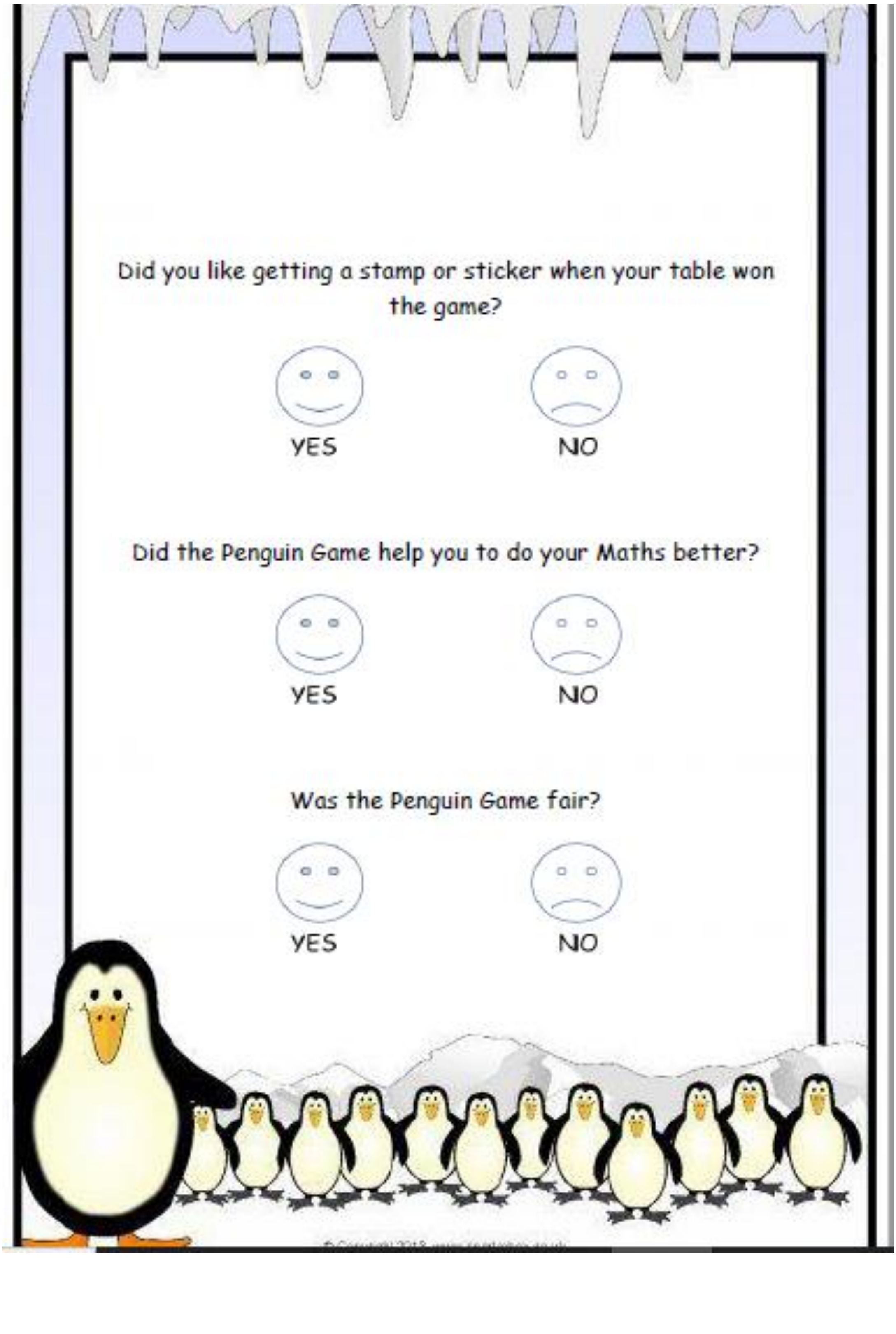


Image downloaded from cool2bekids.com, circles added for awarding points in Microsoft Word.

Appendix BB: CIRP Completed by Ms. Ellis' Senior Infants Class, Chapter 5

Image border copyright: sparklebox.co.uk (free to copy, distribute and display materials given acknowledgement of source)





Did you like getting a stamp or sticker when your table won the game?

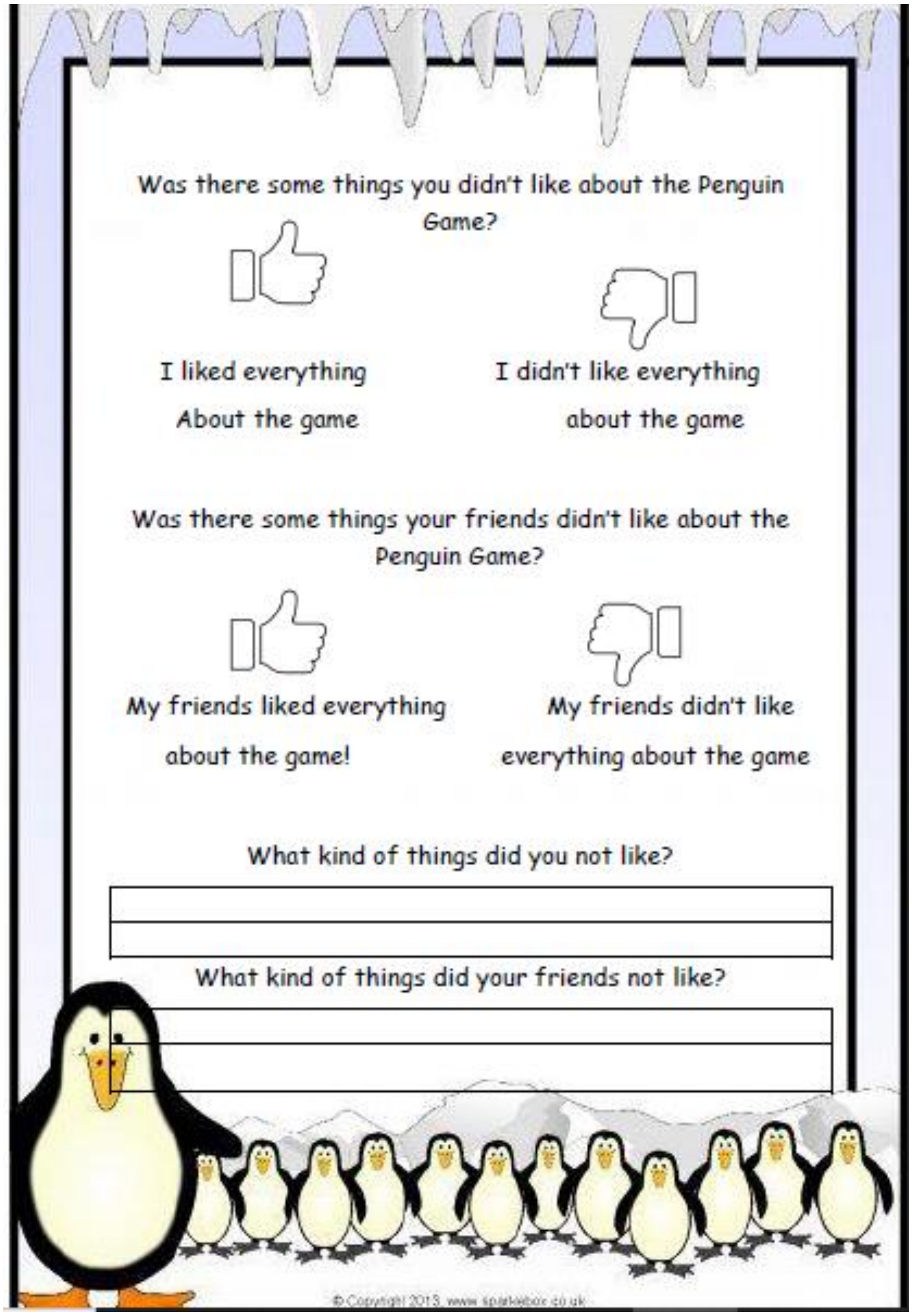
YES NO

Did the Penguin Game help you to do your Maths better?

YES NO

Was the Penguin Game fair?

YES NO



Was there some things you didn't like about the Penguin Game?

I liked everything About the game

I didn't like everything about the game

Was there some things your friends didn't like about the Penguin Game?

My friends liked everything about the game!

My friends didn't like everything about the game

What kind of things did you not like?

What kind of things did your friends not like?

© Copyright 2013 www.sparklebox.co.uk

Reference: Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in School Psychology* (pp. 251–288).

With modifications adapted from 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. Further modifications made to suit young, senior infant population