2019

Evaluating feedback during the Step It Up! game to increase physical activity exhibited by elementary school students during recess

Carla Burji
University of the Pacific, burji.carla@gmail.com

EVALUATING FEEDBACK DURING THE STEP IT UP! GAME TO INCREASE PHYSICAL ACTIVITY EXHIBITED BY ELEMENTARY SCHOOL STUDENTS DURING RECESS

by

Carla Burji

A Thesis Submitted to the

Graduate School

In Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

College of the Pacific
Behavioral Psychology

University of the Pacific
Stockton, California

2019

EVALUATING FEEDBACK DURING THE STEP IT UP! GAME TO INCREASE PHYSICAL ACTIVITY EXHIBITED BY ELEMENTARY SCHOOL STUDENTS DURING RECESS

by

Carla Burji

APPROVED BY:

Thesis Advisor: Matthew Normand, Ph.D.

Committee Member: Corey Stocco, Ph.D.

Committee Member: Katie Wiskow, Ph.D.

Department Chair: Scott Jensen, Ph.D.

Dean of Graduate School: Thomas Naehr, Ph.D.

EVALUATING FEEDBACK DURING THE STEP IT UP! GAME TO INCREASE PHYSICAL ACTIVITY EXHIBITED BY ELEMENTARY SCHOOL STUDENTS DURING RECESS

by

Carla Burji

DEDICATION

This thesis is dedicated to my parents, Nasser and Anabelle Burji, for all their love and support throughout my journey. Thank you for teaching me the value of education and supporting all my decisions.

ACKNOWLEDGEMENTS

I would like to thank Dr. Matthew Normand for his patience, guidance, and valuable feedback throughout this program. I would also like to thank my committee members, Dr. Corey Stocco and Dr. Katie Wiskow, for providing me with the feedback needed to successfully complete my thesis. Most importantly, I would like to thank Alex McCurdy for being exceptionally supportive throughout this process. Lastly, I would like to thank my lab mates, Kelly Roughgarden and Jesse Bober for helping me throughout this process! I could not have done this without the lot of you. Thank you, greatly.

Evaluating Feedback During the Step it UP! Game to Increase Physical Activity Exhibited by Elementary School Students During Recess

Abstract

by Carla Burji

University of the Pacific
2019

Children who are physically inactive are more likely to suffer numerous health complications such as obesity, cardiovascular disease, and diabetes. The Good Behavior Game (GBG) has been identified as an effective method for decreasing inappropriate behaviors and increasing appropriate behaviors in a variety of settings; however, few studies have used the GBG to increase physical activity. Furthermore, no previous research has evaluated the effects of feedback while playing the GBG to increase physical activity. Some research suggests certain characteristics of feedback tend to produce consistent changes in behaviors such as providing feedback immediately and privately. The purpose of the current study was to extend previous research and evaluate whether the addition of feedback to the Step it UP! Game, a modified version of the GBG has an impact on children's step counts. A reversal (ABCAB) design was used to examine the number of steps students took during baseline (A), the Step it UP! Game (B), and the Step it UP! Game with feedback (C). Sealed pedometers were distributed to 21 students from a fifth-grade general education classroom. The results of this study suggest that the Step it UP! Game with feedback did not enhance the number of steps taken during recess. Additionally, the Step it UP! Game (with and without feedback) did result in a slight increase in mean steps per min during recess initially but, these steps did not maintain overtime.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

Chapter 1: Introduction

Physical inactivity is one of the leading risk factors for global mortality (World Health Organization [WHO], 2018) and can result in various health complications overtime such as obesity, cardiovascular disease, cancer, and diabetes (Blair & Brodney, 1999; Myers et al., 2015). Physical activity (PA) plays an important role in children's lives because it helps increase cardiometabolic health as well as musculoskeletal health and fitness (Janssen & LeBlanc, 2010; Väistö et al., 2017). Introducing physical activity to children can lead to many long-term health benefits including weight control, decreased risk of some diseases (e.g., cancer, cardiovascular disease, and type 2 diabetes), as well as increased muscle density, cardiorespiratory health, bone strength, improved mental health, and life expectancy (Center for Disease Control and Prevention [CDC], 2017; Reiner, Niermann, Jekauc, & Woll, 2013; Warburton & Bredin, 2017). To achieve these health benefits, one recommendation is for children and adolescents to engage in at least 60 min or more of moderate-to-vigorous physical activity (MVPA) daily such as running, jumping, and climbing (Brown et al., 2008; CDC, 2018; WHO, 2018). Unfortunately, few children and adolescents engage in the recommended MVPA levels (National Association for Sport and Physical Education, 2016).

Children spend a large portion of their day at school (National Center for Education Statistics [NCES], 2015). Early intervention in elementary school is one strategy to increase children's PA levels (Arundell et al., 2013). Recess periods provide an opportunity for children to engage in PA and unstructured free play (CDC, 2018; NCES, 2005); however, research suggests that children's PA levels are highest during the first 3 min of recess and tend to

decrease as recess progresses (McKenzie et al., 1997; Sallis, Patterson, McKenzie, & Nader, 1988). Therefore, increasing the amount of time spent engaged in PA during recess seems a viable intervention target.

Most school-based interventions during recess periods use antecedent-based strategies such as manipulating the environment (e.g., adding playground markings and/or equipment, Escalante et al., 2014; Hannon & Brown, 2008), providing physical education advice (e.g., Planet Health, Gortmaker et al., 1999), or introducing structured recess (Howe, Freedson, Alhassan, Feldman, & Osganian, 2012). These interventions have been shown to produce moderate or brief changes in behavior, whereas some research, albeit outside of the realm of PA, suggests that persistent change is more likely to occur when consequence-based interventions are implemented (e.g., Roane, Ringdahl, & Falcomata, 2015). Therefore, incorporating a consequence-based intervention that is shown to reliably produce increases in PA during recess would be more useful than most typical school-based interventions, overall.

One way of incorporating consequence-based interventions is through the use of group contingencies. Group contingencies are a common intervention strategy used in school settings to decrease inappropriate behaviors such as elopement and disruption in the classroom (Thorne & Kamps, 2008), and risk-taking behavior on the playground (Heck, Collins, & Peterson, 2001). Furthermore, group contingencies have been used to increase appropriate behaviors such as academic performance (e.g., Greenwood, Hops, Delquadri, & Guild, 1974), prosocial behaviors (e.g., Cashwell, Skinner, & Smith, 2001), and physical activity levels (e.g., Kuhl, Rudrud, Witts, & Schulze, 2015). Group contingencies can either be dependent, independent, or interdependent (Cooper, Heron, & Heward, 2007). Dependent group contingencies occur when a consequence is given to a whole group based on the target behavior(s) of specific individuals of the group

(Cooper et al., 2007; Litow & Pumroy, 1975). Independent group contingencies occur when a consequence is provided only to the individuals in the group who have exhibited the target behavior (Cooper et al., 2007). In classroom settings, interdependent group contingencies are usually applied so that the entire class has the opportunity to contact reinforcement when everyone exhibits the target behavior (Cooper et al., 2007).

The Good Behavior Game (GBG; Barrish, Saunders, & Wolf, 1969) is an interdependent group contingency typically applied in classroom settings that has been shown to be particularly effective at decreasing inappropriate behaviors (e.g., out-of-seat, talking-out). The intervention contains four core components: instruction, reinforcement, teams, and feedback. During the typical GBG, the teacher evenly distributes the students across two teams (team component). The teacher then explains the rules of the game (e.g., instruction component; no talking out of turn, sitting down appropriately) and points are given for breaking any of the rules (e.g., feedback component; out-of-seat, talking-out), which the teacher records on a board that is visible to all students (feedback component). Both teams have the opportunity to win if the points are below a specific criterion, regardless of which team has fewer points (reinforcement component). For example, if one team has three points and the other has four points, both teams win because they are below five points. However, if both teams exceed the criterion (e.g., five points), the team with the fewest points wins the game and everyone on the winning team receives a prize (e.g., reinforcement component; line up first for lunch, stickers).

Since the publication of Barrish et al. (1969), many studies have replicated and extended the GBG across various populations including individuals diagnosed with fetal alcohol syndrome (e.g., Donaldson, Matter, & Wiskow, 2018), high school children (e.g., Flower, McKenna, Muething, Bryant, & Bryant, 2014), and preschool children (e.g., Lynne et al., 2017).

Additionally, the GBG has been applied in different settings such as the school playground (Galbraith & Normand, 2017), physical education class (Jung, Suroto, Fukugasako, & Takahashi, 2005), and the school cafeteria (McCurdy, Lannie, & Barnabas, 2009). Even so, the acceptability of the GBG has shown that it is a useful and easy intervention to implement (Donaldson, Vollmer, Krous, Downs, & Berard, 2011; Flower, McKenna, Bunuan, Muething, & Vega, 2014; Kleinman & Saigh, 2011).

More recently, the GBG has been used to target PA (Galbraith & Normand, 2017; Jung et al., 2005; Normand & Burji, in press). Jung et al. (2005) used a multiple-baseline design across student target behaviors (i.e., waiting time, off-task, motor activity) to examine the effects of a modified version of the GBG during physical education class. One group experienced the GBG, while a control group was provided verbal reminders to follow directions in a timely manner, remain on-task, and engage in motor activity. Jung et al. found that using the GBG reduced student waiting time (i.e., directions were followed in a timely manner), reduced off-task behaviors, and increased motor activity, while the behaviors exhibited by the control group students changed minimally compared to baseline. The results from Jung et al. suggest that the GBG can be used to increase PA. However, these conclusions must be tempered. Jung et al. did not clearly define motor activity and the data collected might not provide an accurate measure for the target behaviors. Motor activity was defined as student engagement with the following activities during physical education class: skill practice, drills, scrimmages, games, fitness activities, warm-up, and cool-down. Additionally, the study used a 10-second discontinuous partial interval recording (5 second observe, 5 second record) method where each student rather than the group was observed for 2 min. This could have resulted in missed occurrences of the target behaviors.

Galbraith and Normand (2017) addressed the limitations of Jung et al. (2005) and evaluated the effects of the Step it UP! Game, a modified version of the GBG on PA exhibited by elementary school children during recess. Students were given sealed pedometers to record the number of steps that they took during recess. Pedometers were used due to a variety of reasons: 1) the steps taken can be directly observed, 2) the steps taken are highly correlated across different devices, and 3) pedometers are widely understood by researchers and the general public (Barreira & Schuna Jr., 2018). Moreover, pedometers provide an observable measure of physical activity that is closely related to the activity data reported by heart rate monitors and accelerometers (McNamara, Hudson, & Taylor, 2010). Galbraith and Normand (2017) used an alternating-treatments design to compare traditional recess periods to recess periods during which the Step it UP! Game was played. Galbraith and Normand found that the Step it UP! Game produced higher mean step counts for both teams compared to baseline. Furthermore, an analysis of each student's step counts showed that 14 out of the 20 students took more steps during the Step it UP! Game than in baseline. One limitation of the Galbraith and Normand (2017) study was that, unlike most GBG studies (Barrish et al., 1969; Donaldson et al., 2011; Lynne et al., 2017) that provide feedback in the form of points on a board and statements for rule violations, Galbraith and Normand (2017) did not provide performance feedback. Additionally, some students did not consistently take more steps during the game compared to baseline.

Throughout the GBG, instruction and reinforcement have been shown to be effective, in and of themselves, whereas feedback is not clearly defined (Embry, 2002; Foley, Dozier, & Lessor, 2018; Harris & Sherman, 1973; Medland & Stachnik, 1972). Therefore, feedback during the GBG may not be as effective as it could be in terms of producing behavior change. Feedback during the GBG is typically provided immediately after the teacher observes a team member

break one of the rules specified at the start of the game (Barrish et al., 1969; Donaldson et al., 2011; Pennington & McComas, 2017; Wiskow, Matter, & Donaldson., 2019). In most cases, the teacher places a mark on the board (visual feedback) and states which team violated the rule (verbal feedback). The board is located where all students can see it throughout the game (Barrish et al., 1969; Groves & Austin, 2017; Salend, Reynolds, & Coyle, 1989). Because feedback is present during the GBG, it is important to evaluate the role it has within the GBG, especially because some research suggests that specific forms of feedback are required to produce changes in behavior (e.g., goal setting, private and public feedback, daily feedback; Alvero, Bucklin, & Austin, 2001; Balcazar, Hopkins, Suarez, 1985).

Although the contingency-specifying parts of the rules for the GBG are largely the same whether targeting problem behavior or PA, the type of feedback and setting differs for studies that use the GBG to target PA (Galbraith & Normand, 2017; Jung et al., 2005). Jung et al. (2005) delivered feedback by publicly posting the accumulated points for each team at the end of the session. Conversely, Galbraith and Normand (2017) delivered verbal prompts throughout recess (on a 3-min variable-time schedule) to remind teams to engage in PA, which could act as feedback. However, the prompt was not considered to be feedback as it did not specify anything about the team's performance, nor did it follow the target behavior of PA. Although both studies used some form of performance feedback, it is unclear whether this component had a notable effect on student's behaviors during the GBG. The current study evaluated whether the provision of group feedback (public posting of team performance) and individual feedback (private posting of individual performance) enhanced the effects of the GBG when used to increase the steps taken by elementary students playing on a school playground during recess.

Chapter 2: Methodology

Participants and Setting

Twenty-one participants (12 girls, 9 boys) from one fifth-grade regular education classroom from a public school located in Stockton, CA participated in the study. Participant's ages ranged from 10 to 11 years old. Additionally, the teacher from the classroom participated in the study by completing a social validity questionnaire at the conclusion of the study. All students in the class participated in recess daily; however, only the students who returned signed consent forms were allowed to participate in the study. No additional demographic information was available for the participants. No compensation was provided for participation. The University of the Pacific institutional review board approved all aspects of the study before it began.

The study took place during regularly-scheduled recess periods on the school playground. The school playground consisted of a fixed play structure, paved play area, and grassy play area. However, per school protocol, students were not allowed to access the grassy play area during recess. Students were allowed to bring out playground equipment such as basketballs and dodgeballs to play in the pavement play area.

Materials

The experimenter provided Yamax SW200 Digi-Walker pedometers to the students to record the number of steps taken during recess. These pedometers have been determined to accurately measure steps taken by young children and, hence, to be acceptable for research purposes (Kilanowski, Consalvi, & Epstein, 1999; Mckee, Boreham, Murphy, & Nevill, 2005).

The pedometers were masked, meaning that participants were not allowed to view their steps throughout recess. The main reason the pedometers were masked was to prevent students from accidentally resetting their pedometers during recess. During the experimental sessions, the researchers planned to have students wear colored flag belts (i.e., green or yellow) to identify which team they were playing on however, due to teacher preference, students were not required to wear the flag belts. As an alternative way to indicate on which team each student was playing, a team list was posted near the classroom door. The experimenter used a MotivAider® (Behavioral Dynamics, Inc.) which vibrated on a specified time schedule and served as a prompt to the experimenter to remind students to continue taking more steps throughout recess (see below). The experimenter also used a smartphone device to access a stopwatch application to time the session and a camera application to take pictures of the pedometers at the end of session. Moreover, plastic containers were used to store pedometers and a large white board and Post-it notes were used to depict the number of steps taken.

Response Definition and Measurement

The dependent variable was the mean number of steps per min recorded for each team during recess and was calculated by taking the mean number of steps and dividing by the duration of the session. The mean number of steps was calculated by adding the total number of steps for each student (recorded by the pedometers) on a team and dividing this sum by the total number of students who participated on that team on that specific day. The mean number of steps was used to announce the winning team because the number of students participating each day varied due to a variety of reasons (e.g., absences). Mean steps per minute was calculated because the duration of the recess periods varied from day to day (range, 9 min, 2 sec to 11 min, 44 sec), mainly due to the amount of time it took students to line up in front of the classroom

once recess was over. Throughout the entire study, no students opted out from participating across all sessions. However, students missed sessions on average 1.95 times (range, 0 to 5) across the study for various reasons: they were absent, they were sent to the office for misbehaving, they forgot to wear their pedometer, or they had not yet returned a signed consent form.

Interobserver Agreement

We calculated interobserver agreement (IOA) for 100% of sessions for baseline and the Step it UP! Game with feedback conditions, as well as for 92% of sessions for the Step it UP! Game without feedback. A picture of the pedometers was not taken during session 26 of the Step it UP! Game without feedback condition due to experimenter error. A second observer reviewed pictures of the pedometers and recorded students' step counts using an Excel spreadsheet. The experimenter calculated IOA by dividing the smaller number of steps for all students by the larger number of steps for all students and multiplying the total by 100 to produce a percentage. During all sessions, the experimenter's IOA on average step counts was 100%.

In addition, the experimenter tested all pedometers at the end of every week to ensure they were calibrated. To do this, the experimenter viewed the Yamax SW200 Digi-Walker instructions on how to ensure pedometer functionality. The experimenter first placed the pedometer vertically in her hand and shook it up and down. If a clicking sound was emitted (pendulum located in pedometer), this indicated that the pedometer was functioning properly. To further ensure for correct calibration, the experimenter measured and marked the distance of a room (0.8 meters), counted the number of steps it took to cover the distance (about 28 to 30 steps), and tested each pedometer by placing them on her waist. Pedometers were replaced if

they were not emitting a clicking sound or if they were plus or minus 3 steps from the minimum or the maximum number of steps covered. During the entirety of the study, only one pedometer had to be replaced due to poor calibration.

Procedure

At the beginning of the study, the experimenter obtained permission from the school principal to conduct the study at the school. Then, the experimenter met with the fifth-grade teacher to explain the purpose of the study, gain her consent, and provide the informed consents for her to distribute to the legal guardians of her students. The study lasted 29 days, with sessions typically conducted five times per week. About halfway through the study, there were 11 days with no sessions conducted due to poor weather conditions (e.g., air quality, rainy days) and because of Thanksgiving break.

A reversal (ABCBC) design was initially proposed, with the A phase representing the baseline condition, B phase representing the Step it UP! Game condition, and C phase representing the Step it UP! Game with feedback condition. Due to the minimal changes in step totals observed during the Step it UP! Game with feedback condition (C), the design was changed to an ABCAB arrangement (see Results below).

Before each session, the experimenter distributed the assigned (masked) pedometers to each student while they were seated. Each student wore a pedometer throughout recess, unless they chose not to participate. After distributing the pedometers, the students sat at their desks and waited for their teacher to signal that they were ready to go outside for recess. Recess began at approximately 9:15 a.m. each day and lasted 11 mins, on average (range, 9 to 12 mins). The experimenter started the stopwatch immediately after the first student exited the classroom. If a student requested to leave the playground or no longer wanted to participate in the study, the

experimenter removed the pedometer from the student for the remainder of the session and the students step count was not recorded for that day. Once recess was over, the students formed a line in front of their classroom. When the last student was in line, the experimenter stopped the stopwatch, collected the pedometers, and placed them in a plastic container. The experimenter then inputted the step data into an Excel spreadsheet and took pictures of the pedometers.

Potential reinforcer survey. Prior to the start of the study, the experimenter met with the teacher to identify potentially reinforcing activities and items in their classroom that could be used during the experimental sessions (see Appendix C). The teacher selected three activities for her students to choose from: 5 extra minutes of recess, lunch in the classroom with the teacher, and a homework pass. The three activities were presented to the whole class and students were asked to vote by raising their hands. Students voted on an extra 5 mins of recess (zero votes), lunch in the classroom with the teacher (10 votes), and a homework pass (12 votes). The activity with the most votes (homework pass) was selected as the prize to be used throughout the experimental sessions. However, prior to starting the intervention, the teacher expressed concern that it was not feasible for students to receive a homework pass every day. Because of this concern, the reward was modified so as to alternate between homework passes and lunch in the classroom with the teacher (e.g., on Monday, the team with the most steps received a homework pass; on Tuesday, the team with the most steps received lunch with the teacher). In addition to homework passes and lunch in the classroom with the teacher, the winning team would earn Step it UP! Champ badges which they would later trade for the larger reward (i.e., lunch with the teacher or a homework pass).

Team distribution. Prior to the start of the study, the experimenter and the teacher discussed whether to use a matched-pairs strategy or random assignment (on a daily basis) to

create the teams. The matched-pairs strategy is meant to equally distribute the students across two teams based on their performance and gender. However, there are times that one team might earn rewards more often than another team, which might make it discouraging for those on the losing team. In other words, failure to contact reinforcement might extinguish physical activity during the game, or it might serve as an abolishing operation that reduces the reinforcing value of winning, or, conversely, as an establishing operation that increases the value of winning, depending on individual learning histories. An alternative is to randomize teams each day to ensure that each student has the opportunity to earn a reward.

After reviewing these options with the teacher, the teacher selected to use a matched-pairs strategy to distribute the teams. Although the teacher selected a matched-pairs strategy, this option is less likely to be adopted in a classroom because of the amount of time it would take for a teacher to calculate step data, determine who to pair, and distribute team members equally. Future studies might want to consider the randomization of teams as this option addresses the issue of social validity.

Math assessment. Prior to the start of the study, the experimenter asked the teacher to distribute a short questionnaire (see Appendix D) that assessed the students' understanding of mathematical averages (means). The purpose of this was to determine whether the students understood basic averages, because the average number of steps for each team was displayed and information about the student's overall performance compared to the winning teams average during the Step it UP! Game with feedback condition were provided. Thus, some understanding of averages was vital for students to respond to the independent variable. It is important to note that the teacher indicated to the experimenter that the school curriculum changed and that students are no longer taught averages in fifth grade but, instead, are taught averages in the sixth

grade. The experimenter provided the initial math assessment (pre-test), then provided a short refresher training on averages, then provided the math assessment again (post-test). During the refresher training, the experimenter defined the term "average," showed the students how to calculate averages, and reviewed relative averages (e.g., performing below, above, or at average levels) with them. If the students had a failing score (less than 80%), the experimenter still continued with the intervention and noted (on the data sheet) which students failed. Fifteen students failed the math assessment (see Results below). The math assessment took approximately 45 mins of class time, in total.

Baseline. Prior to the initial baseline session, the experimenter read the following script to the classroom:

> We want to see how much everyone plays during recess. This is called a pedometer <points to pedometer> and it tells us how many steps you take. You wear it on your waist like this <demonstrates how to clip the pedometer>. We want you to wear one during recess. We will give them to you at the start of recess every day and take them back at the end of recess to see how much you played. It will take about 10 minutes, but you can stop whenever you want. If you want to rest or stop playing, please let us know. You will not get into any trouble. If you decide that you do not want to play anymore, you can give us your pedometer. If you have any trouble, please talk to one of us. Does anyone have any questions?

Following the initial set of instructions, students went out for recess and played as they normally would. The experimenter did not interact with the students and no programmed consequences for PA occurred. There was a total of 10 baseline sessions.

Step It UP! Game. Students were divided into two teams according to a matched-pairs strategy using their baseline data. That is, high performers were matched with low performers with an equal distribution of boys and girls on each team. Before the initial Step it UP! Game session, the experimenter explained the rules of the game and stated the following to the classroom:

> Today, we are going to play the Step It UP! Game and each team will have to be as active as they can while playing outside. You will need to wear your pedometer. Once in a while, we will remind everyone to keep on moving. The more you move, the better. At the end of recess, the team that takes the most steps will win the game and get a prize. It will take about 10 minutes, but you can stop whenever you want. At the end of recess, we will collect your pedometers. If you want to rest, or stop playing, please let us know. You will not get into any trouble. If you decide that you do not want to play anymore, you can give us your pedometer. If you have any trouble, please talk to one of us. Does anyone have any questions?

During recess, the experimenter wore a MotivAider® on their waistband. The MotivAider® vibrated on a 3-min variable-time schedule to remind the experimenter to deliver a verbal prompt (within 3 s of the vibration) to the students during recess. The verbal prompt reminded the students to continue engaging in PA (e.g., "Keep on moving if you want your team to win!").

At the end of recess, the experimenter calculated the average number of steps for each team, announced the winning team (e.g., "the yellow team took the most steps today!"), and immediately delivered the Step it UP! Game badges to the team members. The badge was later traded for either a homework pass or access to lunch with the teacher. Both teams had the opportunity to receive a prize if the average step count difference between both teams were 100 steps or fewer. There was a total of 13 Step it UP! Game sessions.

Step it UP! Game with feedback. This condition was similar to the Step it UP! Game condition; however, the experimenter provided performance feedback at the beginning of each session the day following each game (e.g., students received feedback Tuesday morning based on Monday's results). More specifically, the experimenter wrote the average number of steps for each team on a whiteboard and provided individual step counts for each student on Post-it notes. The Post-it note contained a traffic light (i.e., red light, green light, yellow light). The student's number of steps were written beside one of the colors to indicate whether they performed below average (red light), at average (yellow light), or above average (green light) levels in comparison

to the team with the highest average number of steps. Before the initial session, the experimenter read the following script to the classroom:

> Today, we are going to play the Step It UP! Game, but we will also give you information on how much you moved during recess each morning. The average for each team will be posted on this whiteboard <point to whiteboard> and your score will be given to you with these Post-it notes <point to Post-it notes>. The Post-it note will show you whether you performed above, below, or at average levels in comparison to the winning team's average. You will need to wear your pedometer. At the end of recess, the team that takes the most steps will win the game and get a prize. If you decide that you do not want to play anymore, you can give us your pedometer. If you have any trouble, please talk to one of us. Does anyone have any questions?

The announcement of the winning team and distribution of Step it UP! Game badges remained the same as the Step it UP! Game condition (i.e., provided shortly after recess). There was a total of five Step it UP! Game with feedback sessions.

Social validity and student vote. Upon completion of the study, the experimenter returned to the classroom after two days and asked the teacher to complete a questionnaire that included a 5-point Likert-type scale and a written component (see Appendix E). The teacher was asked to indicate how much they agreed with the following statements: 1) The Step it UP! Game by itself increased physical activity, 2) The combination of feedback and the Step it UP! Game increased physical activity, 3) The intervention seemed simple to implement, 4) The study took too long, and 5) I enjoyed the overall experience. A rating of 1 indicated strongly disagree, 2 indicated disagree, 3 indicated neither agree nor disagree, 4 indicated agree, and 5 indicated strongly agree (see Results below).

The written component of the social validity questionnaire asked the teacher the following questions: 1) "Which part of the intervention did you like the most (e.g., Step it UP! Game, Step it UP! Game and feedback)?" 2) "Was there any portion of the experience you

would like to change? If yes, what change(s) do you recommend?" and 3) "We welcome any additional comments/suggestions. Thank you."

On this same day, the experimenter asked students to vote by raising their hands to indicate whether they preferred to have regular recess or play the Step it UP! Game (without feedback). All 21 students were available to vote (see Results below).

Chapter 3: Results

Math Assessment Scores

Only 19 of the 21 participants were present to take a seven-question math assessment. Not all participants were available to take the math assessment as they were absent on the scheduled day. Mean pretest and posttest scores, as well as each student score from pretest to posttest are depicted in Figure 1. The posttest scores were higher (M = 4.84) in comparison to pretest scores (M = 3.89). Visual inspection of Figure 1 suggests that the changes in scores from pretest to posttest varied across participants. Most students improved after the lesson (12 students), some scored the same (four students), and some did worse (four students).

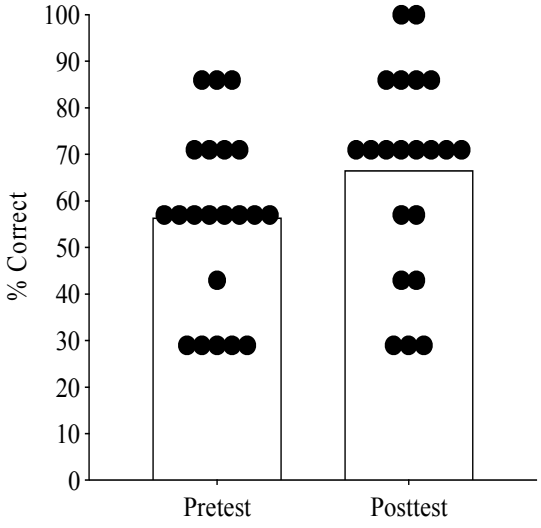

Figure 1: Pre-posttest scatter graph displaying each students score with mean bar graphs.

The number of correct answers by question from pretest to posttest scores also were analyzed. The number of correct answers by question are depicted in Figure 2.

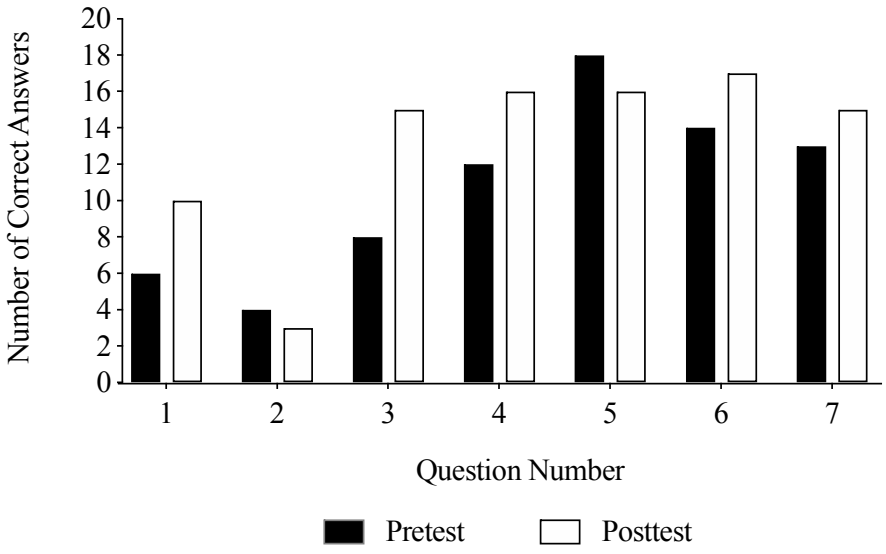

Figure 2: Bar graph depicting number of correct answers from pretest to posttest for each question.

The number of correct answers improved from pretest to posttest for 5 of the 7 questions. Overall, students had the most difficulty calculating the average (questions 1–3). However, students performed better on questions 4–7, which addressed the relative average (e.g., performing above, below, or at average levels). This is important because the relative average was the critical information provided during the Step it UP! Game with feedback.

Step it UP! Game

The top panel of Figure 3 depicts the mean number of steps recorded per min for each condition (i.e., baseline, step it UP! Game, step it UP! Game with feedback) of each team, combined. Mean steps per minute was calculated because the duration of the recess periods

varied from day to day (9 min, 2 sec to 11 min, 44 sec), mainly due to the amount of time it took students to line up in front of the classroom once recess was over. The mean steps per min was calculated by taking the mean number of steps and dividing by the duration of the session.

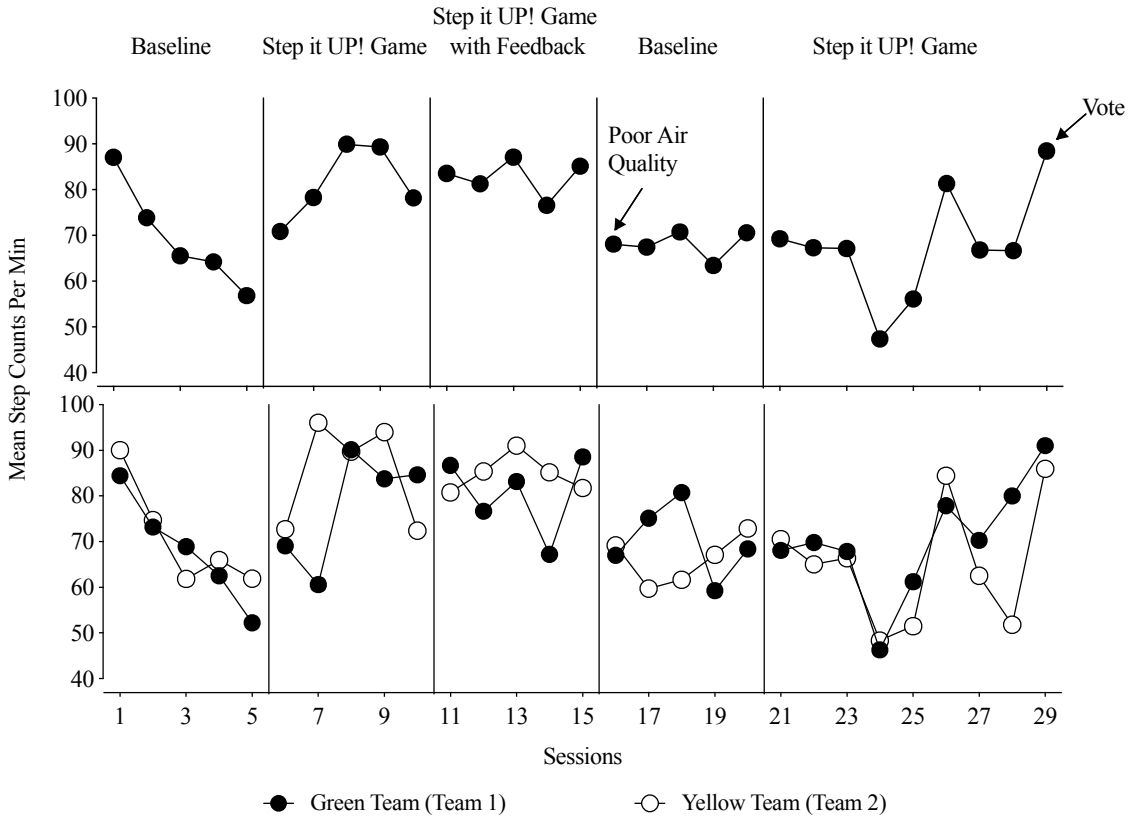

Figure 3: Mean steps per min for both teams (combined), across baseline, Step it UP! Game, and Step it UP! Game with Feedback conditions (top panel). Mean step counts per min for both teams (Green Team and Yellow Team) during each condition (bottom panel).

During the initial baseline, mean number of steps was 68.7 per min (range, 56.9 to 87.1). Visual analysis of the initial baseline reveals a decreasing trend with minimal variability across the five-day period. Higher mean steps per min initially might be due to reactivity to wearing the pedometer. During the initial step it UP! Game condition, the students took 81.3 steps per min,

on average (range, 70.9 to 89.9). The introduction of the Step it UP! Game produced a 12.6 step per min difference in comparison to baseline, suggesting that the game increased physical activity. Figure 3 depicts the mean steps per min during the Step it UP! Game, which gradually increase (more than baseline) and stabilize. Then, the mean steps per min decreased slightly however, continued to be above initial baseline levels. During the Step it UP! Game with feedback condition, the students were taking 82.7 mean steps per min during recess (range, 76.6 to 87.1). Visual analysis reveals that steps during the Step it UP! Game with feedback condition remained high, with levels similar to the Step it UP! Game without feedback. Moreover, the data show low variability with stable responding by the participants. When comparing the Step it UP! Game to the Step it UP! Game with feedback, these data suggest that the Step it UP! Game with feedback did not produce more steps than the Step it UP! Game without feedback. However, students continued to take more steps during the game in comparison to baseline.

During the return to baseline, mean step counts per min were 68.1 (range, 63.4 to 70.8), which was within the initial baseline range. Figure 3 depicts the mean steps per min of the return to baseline, which decrease and stabilize at levels similar to the initial baseline levels (range, 63.4 to 70.8 mean steps per min). Beginning at session 16, there was about a 2-week break due to poor weather conditions (e.g., air quality, rainy days) and because of Thanksgiving break. After the second baseline period, there was a return to the Step it UP! Game (without feedback). The mean step counts per min (65.3; range, 47.4 to 81.3) decreased in comparison to the initial Step it UP! Game condition. Visual analysis reveals two outlier data points (session 24 and 26). Despite the outliers, mean steps per min decreased over time and then, eventually increased before decreasing towards baseline levels. Given this, it might be that the feedback had an effect on performance after all rendering the game without feedback to not be effective (see Discussion

below). A class vote was conducted 2 days later which resulted in students taking an average of 88.5 steps per min during recess—this was considerably more than during the preceding Step it UP! Game (without feedback) condition, and more similar to the initial Step it UP! Game conditions, with and without feedback.

The bottom panel of Figure 3 depicts the mean number of steps recorded per min for each condition (i.e., baseline, step it UP! Game, step it UP! Game with feedback, vote) of each team (green and yellow). For all baseline performances (initial and return to baseline), the green team's mean number of steps were slightly higher (69.2 steps per min; range, 52.3 to 84.4) than the yellow team's mean number of steps (68.5 steps per min; range, 59.8 to 90). Visual analysis of the initial baseline indicates a decreasing trend for both teams. Although there was no "team" composition during baseline, these data are depicted so to analyze the progress of each team throughout the study (e.g., performance levels, equal distribution of wins and losses). During the Step it UP! Game (without feedback), the green team and yellow team were taking about the same mean number of steps, 71.5 steps per min (range, 46.3 to 90.1) and 71.2 steps per min (range, 48.3 to 96), respectively. Visual analysis of the initial Step it UP! Game (without feedback) suggests moderate variability in steps for both teams, with an overall increase in the mean steps per min from the initial baseline. Furthermore, visual analysis of the reintroduction of the Step it UP! Game (without feedback) shows both teams mean steps per min decreasing overtime and then, eventually increasing for the green team, and decreasing for the yellow team. During the Step it UP! Game with feedback condition, the yellow team's mean number of steps were higher (84.8 steps per min; range, 80.7 to 91) than the green team's mean number of steps (80.4 steps per min; 67.2 to 88.6). Both teams continued to take more steps than baseline; however, the yellow team shows more stable responding than the green team (see Figure 3,

bottom panel). During the classroom vote, the green team took more steps (91 steps per min) compared to the yellow team (85.9 steps per min). Taken together, these data suggest that both teams took about the same number of steps across baseline and Step it UP! game sessions. However, the yellow team took more steps during the Step it UP! Game with feedback condition whereas, the green team took more steps during the classroom vote. Overall, both teams took about the same number of steps throughout the study.

Furthermore, there were no obvious patterns associated with a team winning or losing (see Figure 3). That is, if one team won, it did not necessarily mean that the losing team took more or fewer steps that day, or that the winning team would take even more steps the following day. Table 1 presents the number of wins, losses, and ties between the teams, as well as the distribution of wins, losses, and ties by reward (homework pass or lunch with the teacher). There were more tie games (11 ties) with about an equal number of wins for a homework pass (five wins) or lunch with the teacher (six wins). Both teams won about the same number of games (green team won three games and yellow team won four games).

Table 1: Number of Wins, Losses, and Ties per Team

Wins	Green Team	Yellow Team	Both Teams (Tie Game)
Wins for homework pass	2	2	5
Wins for lunch	1	2	6
Total number of wins	3	4	11

Individual step data by each team are depicted in Figure 4. The ratio of steps taken during the Step it UP! Game to steps taken during baseline for each participant was analyzed to compare the intervention effects to baseline. This was done by calculating the average number

of steps a participant took during baseline and dividing this by the average number of steps a participant took during the Step it UP! Game (with and without feedback), to produce a ratio. The higher the ratio, the more steps the participant took during the game compared to baseline. For example, if a participant's ratio was 1.5 and he took 100 steps in baseline, this meant that the participant took approximately 150 steps during the Step it UP! Game.

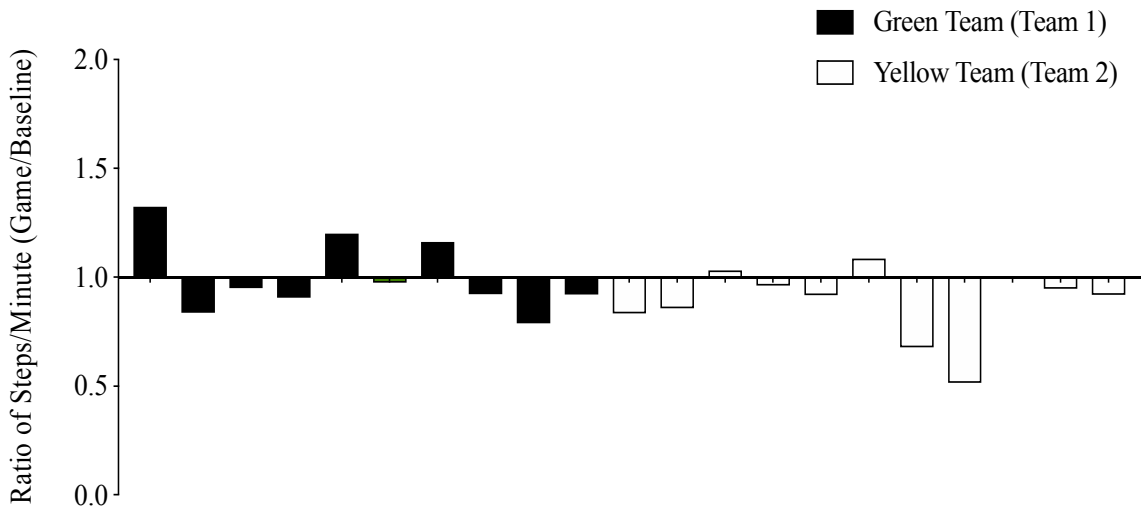

Figure 4: Ratio of steps taken during the Step it UP! Game (with and without feedback) to steps taken during baseline for each participant. Ratios above 1.0 indicate more steps taken during the game.

The individual analysis accounts for all baseline sessions (10 total) and all Step it UP! Game sessions with and without feedback (18 total). The classroom vote was not included in the individual analysis. Only 5 out of the 21 participants took more steps during the Step it UP! Game (with and without feedback) than when they were engaging in regular recess. This was an average increase of 7.5 steps per min (range, 1.2 to 12.9) for those participants, which

corresponds to about 100 to 140 additional steps. This suggests that the Step it UP! Game was only effective at increasing steps more consistently with a small number of participants.

When further analyzing Figure 4, only three participants on the green team took more steps during the Step it UP! Game (with and without feedback) while the other seven participants on the green team seemed to have taken more steps during regular recess (baseline). Moreover, the yellow team only had two participants taking more steps during the Step it UP! Game (with and without feedback) while the remaining nine participants took more steps during recess. Overall, the participants on a given team were not more likely to take more or fewer steps after winning than they were after losing a game (see Figure 3). Furthermore, some students were not present for all sessions, which might have impacted the group's overall performance. That is, students missed on average 1.95 sessions (range, 0 to 5) across the study. More specifically, there were only six students present for all sessions. Four students missed one day (3% of sessions), four students missed two days (7% of sessions), one student missed three days (10% of sessions), four students missed four days (14% of sessions), and two students missed five days (17% of sessions).

Figure 5 depicts the ratio of steps taken during the initial Step it UP! Games (with and without feedback) to steps taken during the initial baseline for each participant. This analysis includes 5 baseline sessions and 10 Step it UP! Game sessions (with and without feedback). During this period, 15 out of the 21 participants took more steps during the Step it UP! Game than when they were engaging in regular recess.

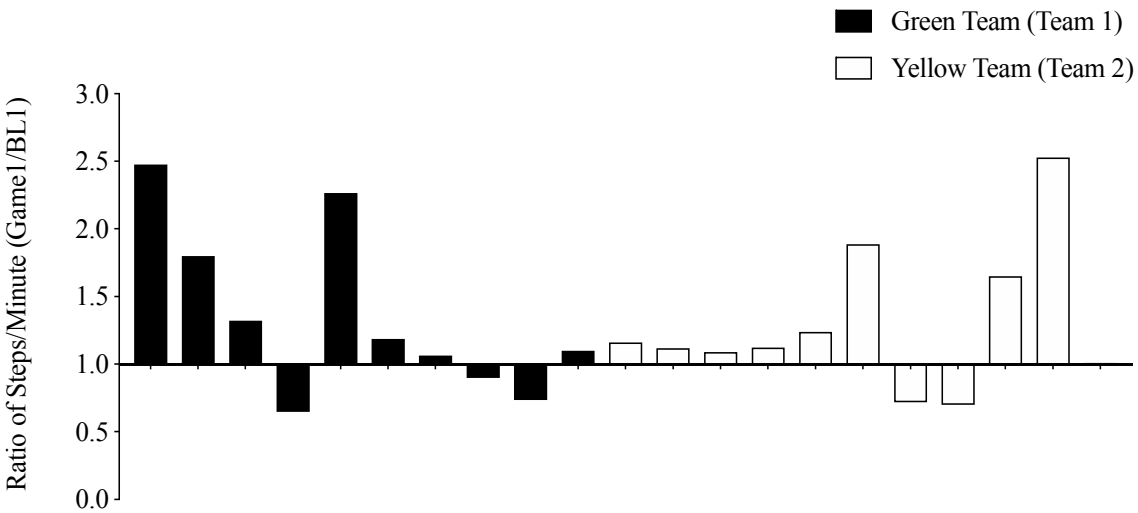

Figure 5: Ratio of steps taken during the initial Step it UP! Games (with and without feedback) to steps taken during the initial baseline for each participant. Ratios above 1.0 indicate more steps taken during the game.

Figure 6 depicts the ratio of steps taken during the second Step it UP! Game (without feedback) to steps taken during the return to baseline for each participant. This analysis includes 5 baseline sessions and 8 Step it UP! Game sessions (without feedback).

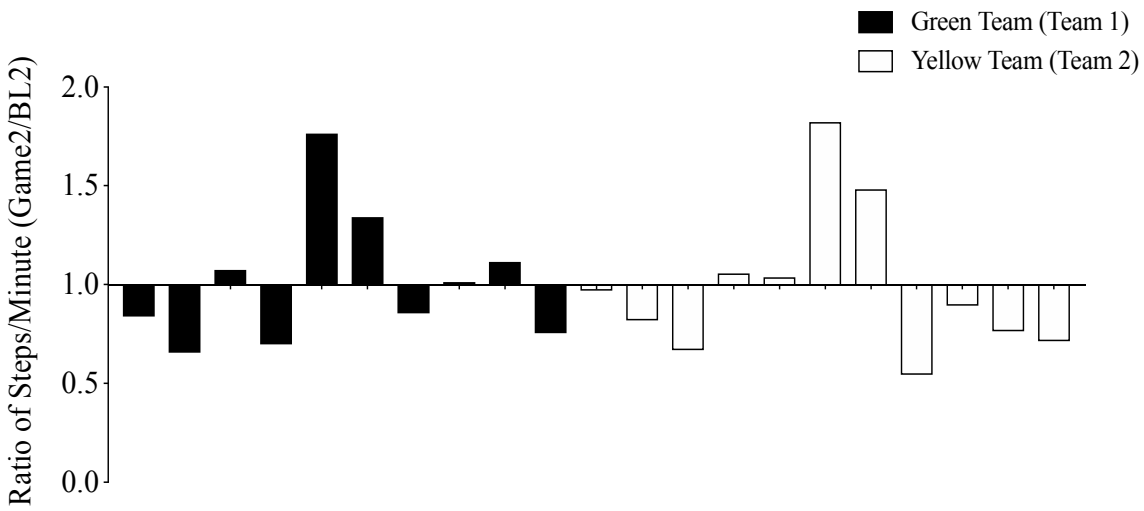

Figure 6: Ratio of steps taken during the second Step it UP! Games (without feedback) to steps taken during the return to baseline for each participant. Ratios above 1.0 indicate more steps taken during the game.

Figure 6 shows that only 9 out of the 21 participants took more steps during the reintroduction of the Step it UP! Game (without feedback) compared to regular recess. That is, 15 out of the 21 participants took more steps during the initial Step it UP! Game (with and without feedback); however, only 9 out of the 21 participants took more steps during the reintroduction of the Step it UP! Game (without feedback).

Social Validity and Student Vote

For the 5-point Likert-type scale, the teacher responded to the first, second, and fifth statement with a "5," indicating the effectiveness and overall experience of the study (see Appendix E). The teacher responded to the third and fourth statement with a "4," indicating that the study took too long but seemed simple to implement. For question 4, the teacher stated she agreed that the study took too long.

During the written component, the teacher answered accordingly: for question one, the teacher responded with "Step it UP! Game with feedback" being the part of the intervention she

liked the most. For question two, the teacher responded with "I would have liked to switch up the prizes half way through. It was a lot of homework passes!" For question three, the teacher stated, "Very cool study. It would be interesting to tie the results or study in an academic sense. I appreciate the organization and the little disruption times. The students very much enjoyed it and created a great relationship tool between the students and I for the lunch prize. Also, they loved the badges and telling staff about the study!" Altogether, the teacher indicated overall agreement to the study.

On the same day as the administration of the social validity questionnaire, the experimenter asked students to vote by raising their hands to indicate whether they preferred to have regular recess or play the Step it UP! Game (without feedback). All 21 students were available to vote, with 19 of the 21 students voting to play the Step it UP! Game. Shortly after voting, the students went outside for recess and played the game.

Chapter 4: Discussion

The current study evaluated whether providing group feedback (public posting of team performance) and individual feedback (private posting of individual performance) enhanced the effects of the GBG when used to increase the steps taken by elementary students playing on a school playground during recess. Previous researchers have used the Step it UP! Game to increase physical activity during recess and physical education (PE) classes (Galbraith & Normand, 2017; Normand & Burji, in press). Both studies have shown that the Step it UP! Game is effective at increasing the number of steps students have taken during recess and PE. The current study utilized the Step it UP! Game procedures and also included a more specific feedback component that most GBG studies use (e.g., Barrish et al., 1969; Wiskow et al., 2019).

Overall, the results of the study show mean step counts increased for both teams when the Step it UP! Game (with and without feedback) was first introduced. That is, there was an average 13.3 step per min increase from the initial baseline to the introduction of the Step it UP! Game (with and without feedback) which corresponds to a 146 step increase during an average 11 min recess period (range, 9 to 12). This would be about an 800 step increase per participant over a one-hour period which might be considered a meaningful change for physical activity (Barreria et al., 2015). However, these results did not maintain during the return to the Step it UP! Game condition (without feedback), which appeared more like the baseline conditions (i.e., 65.2 mean steps per min). Moreover, when the opportunity to participate in the Step it UP! Game (without feedback) versus recess was provided, more steps were observed for only 5 of the 21 participants suggesting that the game with and without feedback was only effective for very

few students (see Figure 4). Furthermore, the addition of feedback did not increase the average number of steps taken during recess (i.e., average of 82.7 steps per min with feedback compared to an average of 81.3 steps per min without feedback).

Given the number of steps taken per min during baseline, it is possible that there was a ceiling effect seen during the interventions; however, previous research has demonstrated that children ages, 9–11 years old typically take 116.5 steps per min when engaged in moderate physical activity and the number of steps tends to decrease as children age (Tudor-Locke et al., 2018). Furthermore, when children (ages, 9–11 years) engage in vigorous physical activity, the steps taken often increase to around 142.7 steps per min (Tudor-Locke et al., 2018). Given these data, it seems that participants in the current study could have taken more steps.

Because the steps-per-min increase did not persist during the return to the Step it UP! Game, it is possible that previously receiving feedback might be a motivating operation that altered the function of consequences for physical activity in the future (Laraway, Syncerski, Michael & Poling, 2003; Michael, 1982, 1993). In other words, the performance feedback that is provided every morning before recess might be an antecedent; specifically, it might be a motivating operation for later activity. The consequences received after playing the Step it UP! Game (with feedback) include the announcement of the winning team (e.g., yellow team wins today) and the provision of rewards (i.e., badges, lunch with the teacher, homework passes). The value of performance feedback for the participants might have been established as an effective reinforcer following the announcement of the winning team and the provision of rewards. The behavior-altering effect might have been that the participants who received rewards after the announcement of the winning team and Post-it notes the following day would be more likely to display behaviors that are associated with winning in the past, such as referring to their Post-it

note to determine whether they need to perform better, and increasing their physical activity levels during recess. Removing the feedback might have functioned as an abolishing operation, reducing the effectiveness of the reinforcers (e.g., badges, lunch with teacher, homework passes) and abating behavior relevant to those reinforcers (e.g., engaging in physical activity).

Despite the minimal change in steps throughout the study, the teacher expressed satisfaction in the effectiveness, feasibility, and experience of the study with the addition of feedback being her most preferred component. Although the teacher expressed feedback as a preferred component, the experimenter spent approximately 30 mins to prepare the feedback for the participants each session. In the long run, the preparation of feedback via post-it notes, and a whiteboard might not be feasible for a teacher to adopt (e.g., consider using technology). The teacher also expressed some concerns about the length of the study (29 days) and the number of homework passes that were used as rewards (10 days of homework passes for at least half the class). Although the length of the study was dependent on the data obtained, it might be important to determine how long a teacher would be willing to adopt and use this intervention on their own.

Additionally, it is unclear whether the teacher's concerns about the length of the study were about the study as a whole taking too long (29 days) or whether the Step it UP! Game portion was too long (18 days). If the teacher's primary concern for the length of the study was about the Step it UP! Game, it suggests that the intervention is not likely to be adopted. In future studies, it will be important to determine which aspects (e.g., length of game component, length of time in classroom) of the Step it UP! Game are acceptable, and to see if teachers actually adopt the Game in their classrooms. Future studies also should assess what type of rewards are most acceptable for the class and teacher. Still, 19 of the 21 participants voted that they wanted

to play the Step it UP! Game (without feedback) instead of regular recess, suggesting that they preferred the game over typical recess periods even without a more extensive preference or reinforcer assessment.

During the Step it UP! Game with feedback condition, participants took about the same mean steps per min as they did during the Step it UP! Game without feedback suggesting that feedback did not have an effect on the average number of steps students took throughout recess. As mentioned previously, feedback during the standard GBG is typically provided throughout the session immediately after the teacher observes a team member break one of the rules specified at the start of the game (Barrish et al., 1969; Donaldson et al., 2011; Pennington & McComas, 2017; Wiskow et al., 2019). During our Step it UP! Game with feedback condition, feedback from the previous day was provided every morning before recess period. The experimenter used feedback characteristics that research suggested were more likely to produce greater increases in performance (Alvero et al., 2001; Balcazar et al., 1995). These characteristics included the source of feedback (i.e., delivered by experimenter), frequency of feedback (i.e., provided daily), and privacy of feedback (i.e., private versus public). During the study, feedback was provided to each participant privately using post-it notes, which indicated the number of steps the participant took the previous day and whether they performed at, above, or below average in comparison to the winning team's average number of steps. Feedback was also provided to the group on a whiteboard indicating each team's average number of steps from the previous day. Although feedback was delivered daily, the feedback was delivered to participants the next day, contributing to a large delay. The delay in feedback might have decreased the effectiveness of the feedback provided (Alvero et al., 2001; Balcazar et al., 1995).

Furthermore, the feedback provided during the Step it UP! Game was not occurring immediately or throughout the session like the standard GBG (Barrish et al., 1969). An analysis of how feedback might have functioned throughout the current study is warranted, and future research should analyze the behavioral principles involved with performance feedback (DiGennaro et al., 2016). Feedback presented prior to a session might act as a discriminative stimulus (S^D). Providing the feedback before the session might evoke the target behavior (e.g., steps) because in the past and under similar circumstances behaving a certain way produced a consequence when the S^D (feedback statement) was present.

In the present study, the Post-it notes, provided to the participants during the Step it UP! Game might have functioned as an S^D or S-delta. The S^D, in this case, is the Post-it note, which indicated whether a participant performed below average (red), at average (yellow), or above average (green). The presence of the Post-it note might have signaled the availability of reinforcement (e.g., badges, homework passes) for increased physical activity and thereby evoked physical activity similar to the previous day. On the other hand, when the Post-it notes were no longer provided (S-delta), this might have given an unclear signal to participants whether reinforcement was available or not, and so participants might have engaged in other behaviors instead (e.g., more gossiping, increased sitting and eating, decreased physical activity). For example, after the Step it UP! Game and feedback condition concluded, baseline was reintroduced. When baseline was reintroduced, participants were no longer receiving feedback or receiving reinforcement (i.e., rewards). Thus, the absence of the Post-it notes coincided with the absence of reinforcement, potentially causing a reduction in physical activity and increase in other behaviors (e.g., gossip).

While attempts were made to provide feedback in a similar manner as previous GBG research (e.g., Barrish et al., 1969; Wiskow et al., 2019), one limitation of the study was that feedback was still not provided immediately throughout the session. Future research might want to examine how to deliver feedback immediately or throughout the session (e.g., use of high-tech devices). One suggestion is to signal to participants throughout the study to regularly view their pedometers. For example, ringing a bell every 2 mins to signal to participants to open their pedometers and observe how many steps they have taken so far. Another suggestion is to consider modifying the verbal prompts provided throughout recess. The experimenter provided verbal prompts to the participants throughout recess. These verbal prompts were reminders to participants to continue engaging in physical activity (e.g., "Keep moving if you want your team to win!" "Get those steps in!"). Although these prompts were provided throughout recess, the prompts were not considered to be feedback as it did not specify anything about the team's performance, nor did it follow the target behavior of PA. Thus, future research might want to consider providing specific prompts contingent on PA (e.g., "Great job moving around!") which might also be a more feasible alternative to providing feedback via post-it notes and a whiteboard.

In addition to the feedback being provided, another limitation of the study is that the participant's might not have had a strong understanding of mathematical averages. Indeed, participants performed better on questions 4–7 of the math assessment, which addressed the relative average. However, there were at least three participants who did not get those answers correct and an additional two participants who did not take the math assessment due to absences. Given the minimal understanding for averages, participants might not have been able to clearly understand the group feedback provided to them, let alone, compare their individual feedback to

the group feedback. As such, this lack of understanding of averages might have changed the participant's performance (e.g., increase, decrease). One way to adjust for this lack of understanding for averages is to provide a refresher training on averages after the first distribution of feedback. For example, the experimenter could create a lesson for the students by using the feedback the students received. Future research might also consider working with a group of participants' who have a better understanding of averages (e.g., older age group).

Another limitation to the study might be the potential reinforcers used throughout the study. While a student vote was conducted to determine the reward to be used throughout the study, the final decision was ultimately made by the teacher to alternate between homework passes (12 votes) and lunch in the classroom with the teacher (10 votes). The participants of the winning team were provided Step it UP! Champ badges that were traded in for homework passes or lunch in the classroom with the teacher. Because of the alternation of rewards, one reward might not have functioned as a reinforcer for some participants compared to the other reward.

For example, 12 students preferred winning homework passes over spending lunch in the classroom with the teacher, while the other 10 students preferred spending lunch with the teacher rather than winning homework passes. This might have influenced behavior in the sense that students who did not want homework passes or lunch with the teacher did not take as many steps during recess, while those students who wanted a specific reinforcer (i.e., homework pass or lunch with teacher) took more steps. Future research should gather data on each individual's vote and conduct an individual analysis on the steps taken throughout recess based on the reinforcers provided. Additionally, it is unclear whether students wanted other rewards. More specifically, most students voted for one item even though they were told they could vote for multiple items. Some students might have voted only for their most-preferred reward rather than

voting for all rewards that they would like. For example, some students might have wanted extra recess if it was available but preferred homework passes and, instead, voted only for the homework passes.

Furthermore, preference or value of a reward might have changed over time so that students who initially preferred a homework pass might find lunch with the teacher more reinforcing over time. Additionally, rewards were alternated daily (e.g., homework passes, lunch with teacher, homework passes, lunch with teacher) with the reward for the day being divulged only after announcing the winning team (e.g., "yellow team won today! You'll be getting homework passes"). However, it became predictable to students what reward they could earn each day. Randomizing when the rewards would be provided so not to be predictable might be an important future step. Additionally, conducting potential reinforcer surveys or using a rotating reinforcer throughout the study might also be an important future step to ensure that participants are getting access to a variety of potentially reinforcing rewards and to prevent satiation of any particular reinforcing item or activity from occurring. It might be beneficial to conduct a vote on a daily basis, or in the case that more than one reward is available, have the students on the winning team select their preferred reward.

During the game, a tie contingency was implemented so that if one team was within 100 mean steps of the other team, both teams could receive a reward. Of the 18 Step it UP! games, participants tied 11 times (61% of sessions). Previous Step it UP! Game studies (Galbraith & Normand, 2017; Normand & Burji, in press) rarely had ties occur. This might be due to having a lower tie criterion (i.e., 50 mean steps) in previous studies. It is possible that the tie criterion might have reduced the effectiveness of the game due to the lack of "competition" involved. Kohn (1992) describes competition as one in which there are winners and losers. The number of

ties were especially prevalent during the reintroduction of the Step it UP! Game (without feedback) condition with ties occurring 75% of the time. Because both teams were contacting reinforcement often, the motivation for participants to take more steps throughout recess might have declined over time.

In addition to the tie criterion, future research should focus on setting a minimum step criterion based on baseline data. For example, each team might have to take at least 900 steps, on average, with the team taking the most steps winning a reward. Another reason why ties occurred so often might have been due to a peer-modeling component occurring during recess. That is, participants' behaviors might not have been entirely independent. Participants might have been paying attention to the environmental events occurring around them, which in turn, could have influenced their behavior. For example, if one group of individuals was especially active, another group nearby might also engage in more physical activity to increase the likelihood of receiving a reward.

When returning back to baseline, a number of sessions were cancelled due to poor weather conditions (e.g., air quality, rainy days) and because of Thanksgiving break. The break between sessions might have resulted in performance changes. More specifically, there were 10 days of no sessions conducted between session 16 and 17, and one day of no session conducted between session 17 and 18 (See Figure 3). Upon returning back to sessions, outdoor conditions were wet and cold due to the rain. Temperatures before returning back to baseline averaged 58 degrees Fahrenheit (range, 48 to 64 degrees) whereas, temperatures upon returning back to baseline averaged 49 degrees Fahrenheit (range, 40 to 56 degrees). Poor weather conditions might have prevented participants from playing in certain areas of the playground (e.g., fixed play structure), possibly leading to lower mean steps recorded. However, a weather analysis was

conducted, and no patterns that might have impacted participants performance during recess were detected. Because weather can be unpredictable at times, future studies might look into alternatives for increasing physical activity that does not include being outdoors (e.g., indoor recess). That is, if indoor recess is provided, examining in what ways physical activity could be introduced indoors (e.g., creating the cafeteria into an open space, using gymnasium space).

In addition to the above limitations, there were some challenges that occurred throughout the study. First, participants stated that some team members were opening their pedometers to view the number of steps taken or removing their pedometers and shaking them. To adjust for this challenge, the experimenter addressed the class that pedometers must be kept closed at all times and that the pedometer would be removed from them and their steps not counted for that day if they were caught cheating. Future research might want to allow participants to look at their pedometers (i.e., unmasked) throughout recess as this might be a more socially valid and significant step. Additionally, future research should set clear expectations on pedometer usage (e.g., steps do not count if you are caught shaking the pedometer). Second, participants did not have access to the grassy play area which might limit the space provided for them to engage in physical activity. This was even more challenging after a rainy day, when play structures were wet, minimizing the play area even more. This could have resulted in a further reduction of the steps taken throughout recess. As mentioned above, future research might want to consider alternatives for increasing physical activity when students do not have access to go outside. Third, the experimenter observed many participants moving their pedometers or placing them in different areas (e.g., pockets, socks, shirt) because of various reasons including, they felt uncomfortable, they kept falling off, or they were wearing dresses. To keep from this happening,

each day the experimenter reminded students to and modeled how to wear their pedometers on their waist (and emphasized the accuracy in steps taken when the pedometer is placed correctly).

A variety of limitations and challenges were mentioned above. However, the reason for the overall reduction occurring after the return to the Step it UP! Game (without feedback) condition is unknown. The current intervention (with and without feedback) did not have an effect on many of the participants. One reason might be a lack of participant interest in the game. Moreover, this is the first Step it UP! Game study to target an older age group (i.e., fifth grade classroom instead of third grade classroom). Some research suggests that elementary school children engage in less physical activity than younger children (e.g., preschool) and that prompts by teachers to engage in physical activity decrease with older age groups (McKenzie et al., 1997). Furthermore, this is one of the first Step it UP! Game studies where the participants were not required to wear flag belts. While this is a cost-effective step, students did not have immediate access to determine who was on each team. The removal of the flag belts could have reduced the "competitive" component of the game despite a team list being posted by the classroom door.

In summary, this study examined whether adding feedback to the Step it UP! Game would increase the number of steps fifth-grade students took during a school recess period beyond what was observed during the Game without feedback. However, the form of feedback provided during this study had no apparent effect on the number of steps participants took throughout recess and so, future research might evaluate alternative ways to provide more effective feedback (e.g., in vivo), if needed. Furthermore, although there was an initial increase in the number of steps taken during the Step it UP! Game, future research should assess ways to

maintain these step increases over time (e.g., feedback provided in vivo, modifying reinforcers, using high-tech devices).

REFERENCES

Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985-1998). *Journal of Organizational Behavior Management, 21*(1), 3–29.

Arundell, L., Ridgers, N.D., Veitch, J., Salmon, J., Hinkley, T., & Timperio, A. (2013). 5-year changes in afterschool physical activity and sedentary behavior. *American Journal of Preventive Medicine, 44,* 605–611. doi:10.1016/j.amepre.2013.01.029

Balcazar, F., Hopkins, B. L., & Suarez, Y. (1985). A critical, objective review of performance feedback. *Journal of Organizational Behavior Management, 7*(3–4), 65-89. doi:10.1300/J075v07n03_05

Barreira, T. V., Schuna Jr, J. M., Mire, E. F., Broyles, S. T., Katzmarzyk, P. T., Johnson, W. D., & Tudor-Locke, C. (2015). Normative steps/day and peak cadence values for united states children and adolescents: National Health and Nutrition Examination Survey 2005-2006. *The Journal of Pediatrics, 166*(1), 139–143. e133. doi:10.1016/j.jpeds.2014.09.014

Barreira, T. V., & Schuna, J. M. (2018). Steps per day and its relationship to energy expenditures. *Medicine & Science in Sports & Exercise, 50*(4), 876. doi:10.1249/MSS.0000000000001509

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*(2), 119–124. doi:10.1901/jaba.1969.2-119

Blair, S. N., & Brodney, S. (1999). Effects of physical inactivity and obesity on morbidity and mortality: current evidence and research issues. *Medicine and Science in Sports and Exercise, 31*, S646-S662.

Brown, W. H., Pfeiffer, K. A., McIver, K. L., Dowda, M., Addy, C. L., & Pate, R. R. (2009). Social and environmental factors associated with preschoolers' nonsedentary physical activity. *Child development, 80*(1), 45–58. doi:10.1111/j.1467-8624.2008.01245.x

Cashwell, T., Skinner, C., & Smith, E. (2001). Increasing second-grade students' reports of peers' prosocial behaviors via direct instruction, group reinforcement, and progress feedback: A replication and extension. *Education and Treatment of Children, 24*(2), 161-175. doi:10.1002/j.2164-4683.1998.tb00405.x

Centers for Disease Control and Prevention [CDC] (2017). Physical inactivity. https://www.cdc.gov/healthcommunication/toolstemplates/entertainmented/tips/PhysicalInactivity.html (accessed on December 26, 2018).

Centers for Disease Control and Prevention [CDC] (2018). Youth physical activity guidelines. https://www.cdc.gov/healthyschools/physicalactivity/guidelines.htm (accessed on December 21, 2018).

Centers for Disease Control and Prevention [CDC] (2018). Recess. https://www.cdc.gov/healthyschools/physicalactivity/recess.htm (accessed on December 26, 2018)

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

DiGennaro Reed, F. D., Henley, A. J., Rueb, S., Crabbs, B., & Giacalone, L. (2016). Discussion of behavioral principles in journal of organizational behavior management: an

update. *Journal of Organizational Behavior Management, 36*(2–3), 202-209. doi:10.1080/01608061.2016.1200938

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. doi:10.1901/jaba.2011.44-605

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. doi:10.1002/jaba.432

Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*(4), 273–297.

Escalante, Y., García-Hermoso, A., Backx, K., & Saavedra, J. M. (2014). Playground designs to increase physical activity levels during school recess: a systematic review. *Health Education & Behavior, 41*(2), 138–144. doi:10.1177/1090198113490725

Flower, A., McKenna, J.W., Bunuan, R.L., Muething, C.S., & Vega, R.V. (2014). Effects of the good behavior game on challenging behaviors in school settings. *Review of Educational Research, 84*(4)*,* 546**–**571. doi:10.3102/0034654314536781

Flower, A., McKenna, J., Muething, C. S., Bryant, D. P., & 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. doi:10.1177/0145445513507574

Foley, E., Dozier, C., & Lessor, A. (2019). Comparison of components of the Good Behavior Game in a preschool classroom. *Journal of Applied Behavior Analysis, 52*(1), 84–104. doi:10.1002/jaba.506

Galbraith, L. A., & Normand, M. P. (2017). Step it up! using the good behavior game to increase physical activity with elementary school students at recess. *Journal of Applied Behavior Analysis, 50*(4), 856–860. doi:10.1002/jaba.402

Greenwood, C. R., Hops, H., Delquadri, J., & Guild, J. (1974). Group contingencies for group consequences in classroom management: A further analysis. *Journal of Applied Behavior Analysis, 7*(3), 413–425. doi:10.1901/jaba.1974.7-413

Gortmaker, S. L., Peterson, K., Wiecha, J., Sobol, A. M., Dixit, S., Fox, M. K., & Laird, N. (1999). Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Archives of Pediatrics & Adolescent Medicine, 153*(4), 409–418. doi:10.1001/archpedi.153.4.409

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. doi:10.1002/jaba.393

Hannon, J., & Brown, B. (2008). Increasing preschoolers' physical activity intensities: An activity-friendly preschool playground intervention. *Preventive Medicine, 46*(6), 532–536. doi:10.1016/j.ypmed.2008.01.006

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. doi:10.1901/jaba.1973.6-405

Heck, A., Collins, J., & Peterson, L. (2001). Decreasing children's risk taking on the playground. *Journal of Applied Behavior Analysis, 34*(3), 349–352. doi:10.1901/jaba.2001.34-349

Howe, C. A., Freedson, P. S., Alhassan, S., Feldman, H. A., & Osganian, S. K. (2012). A recess intervention to promote moderate-to-vigorous physical activity. *Pediatric Obesity, 7*(1), 82–88. doi:10.1111/j.2047-6310.2011.00007.x

Janssen, I., & LeBlanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity, 7*(1), 40.

Jung, J., Suroto, Fukugasako, Y., & Takahashi, T. (2005). The effects of group contingency strategy on students' behaviors in physical education classes in Korea using multiple baseline design. *International Journal of Sport and Health Science,* 3, 46–56. doi: 10.5432/ijshs.3.46

Kilanowski, C. K., Consalvi, A. R., & Epstein, L. H. (1999). Validation of an electronic pedometer for measurement of physical activity in children. *Pediatric Exercise Science, 11*(1), 63–68. doi:10.1123/pes.11.1.63

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. doi:10.1177/0145445510392213

Kuhl, S., Rudrud, E. H., Witts, B. N., & Schulze, K. A. (2015). Classroom-based interdependent group contingencies increase children's physical activity. *Journal of Applied Behavior Analysis, 48*(3), 602–612. doi:10.1002/jaba.219

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*(3), 407–414. doi:10.1901/jaba.2003.36-407

Litow, L., & Pumroy, D.K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis, 8*(3), 341–347. doi:10.1901/jaba.1975.8-341

Lynne, S., Radley, K. C., Dart, E. H., Tingstrom, D. H., Barry, C. T., & Lum, J. D. (2017). Use of a technology‐enhanced version of the good behavior game in an elementary school setting. *Psychology in the Schools, 54*(9), 1049–1063. doi:10.1002/pits.22043

McCurdy, B. L., Lannie, A. L., & Barnabas, E. (2009). Reducing disruptive behavior in an urban school cafeteria: An extension of the good behavior game. *Journal of School Psychology, 47*(1), 39–54. doi:10.1016/j.jsp.2008.09.003

Mckee, D. P., Boreham, C. A., Murphy, M. H., & Nevill, A. M. (2005). Validation of the Digiwalker™ Pedometer for Measuring Physical Activity in Young Children. *Pediatric Exercise Science, 17*(4), 345–352. doi:10.1123/pes.17.4.345

McKenzie, T. L., Sallis, J. F., Elder, J. P., Berry, C. C., Hoy, P. L., Nader, P. R., ... & Broyles, S. L. (1997). Physical activity levels and prompts in young children at recess: a two-year study of a bi-ethnic sample. *Research Quarterly for Exercise and Sport, 68*(3), 195–202. doi:10.1080/02701367.1997.10607998

McNamara, E., Hudson, Z., & Taylor, S. J. (2010). Measuring activity levels of young people: the validity of pedometers. *British Medical Bulletin, 95*(1), 121–137. doi:10.1093/bmb/ldq016

Medland, M. B., & Stachnik, T. J. (1972). Good-behavior game: A replication and systematic analysis. *Journal of Applied Behavior Analysis, 5*(1), 45–51. doi:10.1901/jaba.1972.5-45.

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37*(1), 149–155. doi:10.1901/jeab.1982.37-149

Michael, J. (1993). Establishing operations. *The Behavior Analyst*, *16*(2), 191–206. doi:10.1007/BF03392623

Myers, J., McAuley, P., Lavie, C. J., Despres, J. P., Arena, R., & Kokkinos, P. (2015). Physical activity and cardiorespiratory fitness as major markers of cardiovascular risk: their independent and interwoven importance to health status. *Progress in Cardiovascular Diseases, 57*(4), 306–314. doi:10.1016/j.pcad.2014.09.011

National Association for Sport and Physical Education. (2016). Shape of the Nation report: Status of physical education in the USA. Retrieved from Reston, VA: https://www.shapeamerica.org//advocacy/son/2016/upload/Shape-of-the-Nation-2016_web.pdf (accessed December 21, 2018).

National Center for Education Statistics (2005). State Education Reforms. https://nces.ed.gov/Pubs2006/nutrition/02E.asp (accessed on December 26, 2018).

Normand, M. P., & Burji, C. (in press). Using the Step it UP! Game to increase physical activity during physical-education classes. *Journal of Applied Behavior Analysis*.

Pennington, B., & McComas, J. J. (2017). Effects of the good behavior game across classroom contexts. *Journal of Applied Behavior Analysis, 50*(1), 176–180. doi:10.1002/jaba.357

Reiner, M., Niermann, C., Jekauc, D., & Woll, A. (2013). Long-term health benefits of physical activity–a systematic review of longitudinal studies. *BMC Public Health*, *13*(1), 813. doi:10.1186/1471-2458-13-813

Roane, H. S., Ringdahl, J. E., & Falcomata, T. S. (Eds.). (2015). *Clinical and organizational applications of applied behavior analysis*. Academic Press.

Salend, S., Reynolds, C., & Coyle, E. (1989). Individualizing the good behavior game across type and frequency of behavior with emotionally disturbed adolescents.

Behavior Modification, 13(1), 108–126. doi:10.1177/01454455890131007

Sallis, J. F., Patterson, T. L., McKenzie, T. L., & Nader, P. R. (1988). Family variables and physical activity in preschool children. *Journal of Developmental & Behavioral Pediatrics, 9*(2), 57–61. doi:10.1097/00004703-198804000-00001

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. doi:10.1007/BF03391723

Tudor-Locke, C., Schuna, J. M., Han, H., Aguiar, E. J., Larrivee, S., Hsia, D. S., ... & Johnson, W. D. (2018). Cadence (steps/min) and intensity during ambulation in 6–20 year olds: the CADENCE-kids study. *International Journal of Behavioral Nutrition and Physical Activity, 15*(1), 20. doi:10.1186/s12966-018-0651-y

Väistö, J., Eloranta, A.-M., Viitasalo, A., Tompuri, T., Lintu, N., Karjalainen, P., . . . Lakka, H.-M. (2014). Physical activity and sedentary behaviour in relation to cardiometabolic risk in children: Cross-sectional findings from the Physical Activity and Nutrition in Children (PANIC) Study. *International Journal of Behavioral Nutrition and Physical Activity, 11*(1), 55. doi:10.1186/1479-5868-11-55

Warburton, D. E., & Bredin, S. S. (2017). Health benefits of physical activity: a systematic review of current systematic reviews. *Current Opinion in Cardiology*, *32*(5), 541–556. doi:10.1097/HCO.0000000000000437

Wiskow, K., M., Matter, A., & Donaldson, J. M. (2019). The Good Behavior Game in preschool classrooms: An evaluation of feedback. *Journal of Applied Behavior Analysis, 52*(1), 105–115. doi:10.1002/jaba.500

World Health Organization [WHO] (2018). Physical Activity.

http://www.who.int/mediacentre/factsheets/fs385/en/ (accessed on December 21, 2018)

APPENDIX A: LITERATURE REVIEW

Physical inactivity is the fourth leading risk factor for global mortality (World Health Organization [WHO], 2017) and can result in various health complications such as obesity, cardiovascular disease, and diabetes (Reilly & Kelly, 2011). The WHO (2017) defines physical activity (PA) as, "any bodily movement produced by skeletal muscles that requires energy expenditure." PA plays an important role in children's lives because it is associated with multiple health benefits including: weight control, decreased risk of some diseases such as cancer (i.e., colon, breast), cardiovascular disease, and type 2 diabetes. Other benefits include increased muscle density, cardiorespiratory health, bone strength, improved mental health, life expectancy, balance, and coordination skills (Janssen & LeBlanc, 2010; Reiner, Niermann, Jekauc, & Woll, 2013; Warburton & Bredin, 2017).

In order to achieve the stated health benefits, the WHO (2017) and Center for Disease Control and Prevention (CDC; 2016) recommend that children and adolescents engage in at least 60 min or more of moderate-to-vigorous physical activity (MVPA) daily such as brisk walking, running, jumping, and climbing. However, the National Physical Activity Plan Alliance (2016) released a report on PA for children and adolescents indicating that only 21.6% of individuals engage in the recommended MVPA levels, which reduces their ability to achieve health benefits. In other words, many children and adolescents are not meeting the recommended PA guidelines and instead are engaging in sedentary behaviors (Arundell et al., 2013; Fairclough & Stratton, 2006; Pate, Long, & Heath, 1994; Sallis, Prochaska & Taylor, 2000; Stone, McKenzie, Welk, & Booth, 1998). Therefore, it is important to increase PA levels to counteract the health risks

associated with physical inactivity. Intervening as early as elementary school might be a first step towards increasing PA (Arundell et al., 2013).

School can be an important environment for promoting PA because children spend between 6 and 7 hours at school for approximately 180 days of the year (National Center for Education Statistics [NCES], 2015). Moreover, recess periods during school provide an opportunity for children to engage in PA and unstructured free play (NCES, 2005). For instance, Mota et al. (2005) found that children engaged in about 5% to 40% of the recommended PA levels for the entire day during recess periods. Recess occurs one to three times per day for 15 to 30 min each time in approximately 92% of elementary schools in the United States of America (NCES, 2005). According to the CDC (2017), recess provides the following benefits and changes in behavior: increased PA, continued on-task behavior in the classroom, and reduced disruptive behaviors. Additionally, some research suggests that recess might play a role in improving memory, attention, and concentration as well as, promoting social and emotional development (CDC, 2017; Pellegrini, 1995).

Other research, however, has reported that children's PA levels are highest during the first 3 min of recess and tend to decrease as recess progresses (McKenzie et al., 1997; Sallis, Patterson, McKenzie, & Nader, 1988). This suggests that an intervention that aims to increase PA during the entire recess period might help children meet the CDC's recommended PA guidelines. Due to multiple health benefits associated with recess and the amount of time children spend at school, school-based interventions targeting PA could be an effective and convenient way towards increasing PA, which might also decrease the likelihood of future health complications (Erwin, Ickes, Ahn, & Fedewa, 2014; Hills, Dengel, & Lubans, 2015; Ramstetter, Murray, & Garner, 2010).

Multiple school-based interventions to increase PA have been evaluated (Harris, Kuramoto, Schulzer, & Retallack, 2009; Escalante, Garcia-Hermoso, Backx, & Saavedra, 2014; Van Sluijs, McMinn, & Griffin, 2007). Some interventions have taught physical and health education such as Planet Health, a curriculum aimed to reduce physical inactivity and promote a healthy diet (Gortmaker et al., 1999). Some interventions have altered the existing school play environments to include playground equipment, playground markings, and game equipment (Escalante et al., 2014; Hannon & Brown, 2008). Other interventions have introduced structured recess, which involves planned and supervised activities (Howe, Freedson, Alhassan, Feldman, & Osganian, 2012). And some interventions have implemented multiple components (e.g., health education, physical education, and classroom nutrition; McMurray et al., 2002; Simons-Morton, Parcel, Baranowski, Forthofer, & O'Hara, 1991). Many of these school-based interventions did produce modest increases in PA during a short period of time. However, only a few have yielded socially significant increases or examined whether the PA levels persisted over time. For example, Howe et al. (2012) compared structured recess at one school to free-play recess at another school to determine the effects of recess activities on MVPA for third-grade students over a 9-week period. Results indicated that structured recess produced higher levels of MVPA compared to free-play recess; however, the study was relatively short, which makes it difficult to draw any conclusions about maintenance effects. Furthermore, most of these interventions used antecedent-based strategies, which have been shown to produce moderate or brief changes in behavior, however persistent change is more likely when consequence-based interventions are implemented (e.g., Roane, Ringdahl, & Falcomata, 2015). That is, the literature suggests that antecedent interventions produce temporary changes and those changes

do no persist when not coupled with programmed consequences. Therefore, it might be useful to identify consequence-based interventions that reliably produce increases in PA during recess.

Group Contingencies and the Good Behavior Game

One way of increasing PA is through the use of group contingencies in school settings. Cooper, Heron, and Heward (2007) define group contingencies as ones in which a target behavior is exhibited by either one member of a group, part of a group, or everyone in the group, which then results in a consequence (e.g., reward or removal of a reinforcer). There are three types of group contingencies: dependent, independent, and interdependent. Dependent group contingencies occur when a consequence is given to the whole group based on the target behavior(s) of specific members of a group. Independent group contingencies occur when a consequence is provided only to the individuals in the group who have exhibited the target behavior(s). Interdependent group contingencies provide consequences to the whole group only when everyone in the group exhibits the target behavior(s). A number of studies have demonstrated that group contingencies decrease inappropriate behaviors, including elopement and disruption in the classroom (Thorne & Kamps, 2008), and risk-taking behavior on the playground (Heck, Collins, &, Peterson, 2001). Furthermore, studies have shown that group contingencies increase appropriate behaviors such as academic performance and prosocial behaviors in the classroom (Cashwell, Skinner, & Smith, 2001; Greenwood, Hops, Delquadri, & Guild, 1974), and PA during recess (Kuhl, Rudrud, Witts, & Schulze, 2015).

One version of an interdependent group contingency typically implemented in classrooms, the Good Behavior Game (GBG), has been shown to be particularly effective (Barrish, Saunders, & Wolf, 1969). During the typical GBG, the teacher evenly distributes the students across two teams. In some cases, the teacher evenly distributes the students who exhibit

the most problem behavior across both teams. Then, the teacher explains the rules of the game (e.g., no talking out of turn, sitting down appropriately, being on-task) and points are given for breaking the rules (e.g., out-of-seat, talking-out). The teacher records the points on a board that is visible to all students. Both teams have the opportunity to win if the points are below a specific criterion, regardless of which team has fewer points. For example, if one team has three points and the other has four, both teams win because they are below five points. However, if both teams exceed the criterion, the team with the fewest points wins the game and everyone on the winning team receives a prize (e.g., line up first for lunch, stickers).

Decades of research has demonstrated the effectiveness of the GBG for decreasing problem behaviors and increasing appropriate behaviors (e.g., Barrish et al., 1969; Bunuan, Muething, & Vega, 2014; Donaldson, Fisher, Kahng, 2017; Flower, McKenna, Kellam & Anthony, 1998; Joslyn, Vollmer, & Hernandez, 2014; Wahl, Hawkins, Haydon, Marsicano, & Morrison, 2016; Wiskow, Matter, & Donaldson, 2018). Additionally, the GBG has been modified in multiple ways in order to be applied effectively with various populations and in various settings (Donaldson, Matter, & Wiskow, 2018; Galbraith & Normand, 2017; McCurdy, Lannie, & Barnabas, 2009; Jung, Suroto, Fukugasako, & Takahashi, 2005; Lutzker & White-Blackburn, 1979; Swain, Allard, & Holborn, 1982; Sy, Gratz, & Donaldson, 2016; Wahl et al., 2016). Not only can the GBG be used to change behavior in the present, Embry (2002) emphasized that the GBG could be used as a universal "behavioral vaccine" to prevent the development of inappropriate behavior in the future (e.g., substance use, antisocial behavior, aggressive behavior). For instance, Kellam et al. (2008) assessed the GBG as a universal behavioral vaccine by conducting a longitudinal study in Baltimore beginning in 1985. During this study, the GBG was implemented in first- and second-grade classrooms and was continued

with the same students for 2 years. Schools were randomly assigned to one of three conditions: GBG, curriculum-and-instruction program, and standard program. Kellam et al. found that primary students playing the GBG engaged in fewer inappropriate behaviors compared to students in schools with the curriculum-and-instruction program, and standard program. As part of the longitudinal study, a follow up by Kellam et al. examined the effects of the GBG on young adult outcomes when participants from the 1985 study reached 19 to 21 years of age. This examination found significantly lower levels of problem behaviors (e.g., drug and alcohol use, smoking, violent behaviors) among the GBG intervention population compared with the individuals who experienced curriculum-and-instruction program, and standard program classrooms (Kellam et al., 2008). These results support the idea that the GBG could be used not only to reduce problem behaviors and increase appropriate behaviors in the classroom, but also as a universal behavioral vaccine to prevent problem behaviors later in life.

The acceptability of the GBG amongst teachers and students further increases its value. Students and teachers previously reported that the GBG procedure is an acceptable intervention and a useful tool for reducing behavior problems (Barrish et al., 1969; Flower et al., 2014). For example, the teachers involved in the Barrish et al. (1969) study stated, "it was an easy program to install since it did not change any of the rules or daily activities in the classroom" (p.123). Other studies collected reports from students and teachers to assess social validity of the GBG and found that the GBG was a useful and easy intervention to implement in the school settings (Donaldson, Vollmer, Krous, Downs, & Berard., 2011; Galbraith & Normand, 2017; Kleinman & Saigh, 2011; Lynne et al., 2017). Given the success the GBG has had with multiple populations and settings, the GBG might also be an effective intervention that is favorable to teachers and students for increasing PA.

Physical Activity During the Good Behavior Game

To date, only two studies have examined the effects of GBG on PA (Galbraith & Normand, 2017; Jung et al., 2005). Jung et al. (2005) used a multiple-baseline design across student target behaviors to examine the effects of a modified version of the GBG during physical education class. The study evaluated three target behaviors (waiting time, off-task, motor activity) over a 12-week period. Jung et al. divided the experimental group into four teams with each team assigned a colored vest. The teacher provided rules about the game (e.g., all groups could win, points were given every time a signal went off). Contingent on achieving their behavior goals, the teacher awarded points and delivered rewards to each group (e.g., posting pictures, using a computer during lunchtime) throughout the intervention phase; conversely, points were subtracted for rule breaking. The control group received verbal reminders from the teacher about engaging in appropriate behaviors (i.e., waiting, remaining on-task, and engaging in motor activity), but they did not receive points or rewards.

The results of Jung et al. (2005) indicated that the intervention reduced student waiting time (i.e., directions were followed in a timely manner), off-task behaviors, and increased motor activity, while the behaviors exhibited by the control group changed minimally compared to baseline (Jung et al., 2005). Overall, using the GBG for the experimental group not only decreased inappropriate behaviors (i.e., off-task) but also increased appropriate behaviors (i.e., motor activity, waiting) compared to the control group. The results from Jung et al. study suggest that the GBG can be used to increase PA. However, these conclusions must be tempered because Jung et al. did not clearly define motor activity. Motor activity was defined as student engagement with the following activities during physical education class: skill practice, drills, scrimmages, games, fitness activities, warm-up, and cool-down. Additionally, the data collected

might not have provided an accurate measure of the target behaviors, as data were collected for a portion of the students for a small amount of time, each day. More specifically, data were collected by observing selected students through video recording at approximately 2 min per student. Then, the experimenter coded the target behaviors using a 10-s discontinuous partial-interval recording system.

More recently, Galbraith and Normand (2017) evaluated the effects of a variation of the GBG on PA exhibited by 20 elementary school children during recess. To address the limitations of Jung et al. (2005), Galbraith and Normand used pedometers as a more objective measure of PA, incorporated a within-subject research design, and collected experimenter integrity data. The study took place during recess on an elementary school playground that included pavement, a grassy area, and a fixed play structure. Students were given sealed pedometers to record the number of steps that each student took during recess. Each pedometer was sealed with tape to prevent students from seeing their step totals. The reason for sealed pedometers was to avoid from students accidentally resetting their pedometers. The differences in step counts between baseline and GBG conditions were evaluated using an alternating-treatments design. During baseline, the students were instructed to wear a pedometer during recess and to play as they normally did. Before each GBG session, the experimenter described the rules of the game (e.g., the team with the higher number of steps wins a lottery ticket) and divided students into two teams using colored flag belts. During the GBG, the experimenters reminded the students to engage in PA (e.g., "The more you move, the more you can help your team") according to a 3-min variable-time schedule. At the end of the session, the experimenter collected the pedometers, calculated the mean number of steps for each team, and announced the winning team. The winning team was immediately given a "Step it UP! Champ" badge. At the

end of the day, badges were exchanged for lottery tickets where students could earn prizes from a school-wide raffle. Galbraith and Normand found that the GBG produced higher mean step counts for both teams compared to baseline. Furthermore, an analysis of each student's step counts showed that 14 out of the 20 students took more steps during the GBG than in baseline.

One limitation of the Galbraith and Normand (2017) study was that, unlike most GBG studies (Barrish et al., 1969; Donaldson et al., 2011; Lynne et al., 2017) that provide feedback in the form of points on a board and statements for rule violations, Galbraith and Normand (2017) did not provide performance feedback. Instead, Galbraith and Normand (2017) delivered audible prompts to approximate the feedback that is usually provided during the GBG each time a mark is placed on the board and the prompts delivered were not contingent on physical activity or inactivity. The major difference is that the prompts provided were not the same as the feedback provided during the standard GBG. Additionally, it was difficult for the experimenter to deliver audible prompts to all students because they were in different areas of the playground. Also, some students did not consistently take more steps during the GBG. As such, it would be useful to further the feedback component and determine whether that might produce higher levels of PA.

Performance Feedback

Feedback is a popular intervention component that has been used to increase or decrease a variety of behaviors, including horseback-riding (Kelley & Miltenberger, 2016), coaching (Stokes, Luiselli, Reed, & Fleming, 2010), driving speed (Houten & Nau, 1983), gun safety (Miltenberger et al., 2004), work performance (Newby & Robinson, 1983), and sports performance (Quinn, Miltenberger, Abreu, & Narozanick, 2017). Feedback occurs when information about an individual's performance or particular behavior is provided to them by

other individuals (e.g., supervisors, peers), themselves, or equipment (e.g., pedometers), which can promote behavior change over time (Alvero, Bucklin, & Austin, 2001; Balcazar, Hopkins, Suarez, 1985; Lee, Nyity, & McGill, 1993). In short, the purpose of feedback is to improve future performance. Feedback is often used alone or as an additional component of other interventions due to its practicality, simplicity, and low cost (Alvero et al., 2001; Duncan & Bruwelheide, 1985; Houmanfar, 2013; Prue & Fairbank, 1981). There are a range of characteristics associated with feedback, including the source (e.g., researcher, supervisor), mode of delivery (e.g., oral, written), content (e.g., individual's performance compared to the group's performance), individual receiving feedback (e.g., children, adults), frequency of feedback (e.g., daily, weekly), and privacy (e.g., privately, publicly).

Balcazar et al. (1985) reviewed studies from multiple journals on the consistency of feedback effects by analyzing feedback alone as well as feedback combined with other interventions. Additionally, Balcazar et al. assessed different feedback characteristics (i.e., source, privacy, participants, content, mechanism, frequency) to determine which produced the most consistent effects. They considered the effects of feedback to be consistent when there was a clear increase or decrease in performance compared to baseline and when the change in performance was observed among all participants, settings, or behaviors. In conclusion, Balcazar et al. found that feedback alone does not consistently improve performance, adding behavioral consequences or goals is likely to improve the effectiveness of feedback, and specific characteristics of feedback such as the frequency of delivery and form of delivery are more likely to produce a greater increase in performance than others. These findings suggest that there are particular characteristics of feedback that should be considered when using feedback as part of an intervention.

In a more recent review of the feedback literature, Alvero et al. (2001) reported findings similar to Balcazar et al. (1985). Alvero et al. (2001) found that feedback by itself does not produce consistent effects and that different modes of delivering feedback are more effective (e.g., use of goal setting, private and public feedback, and daily and weekly feedback). However, Alvero et al. (2001) noted some important differences between their review and the initial review by Balcazar et al. (1985). For instance, the number of articles addressing feedback published in the journals reviewed seems to have decreased, and the use of feedback terminology has appeared less frequently. Furthermore, Alvero et al. (2001) examined additional feedback combinations (e.g., feedback alone, feedback and antecedents, feedback and goal setting), as well as subcategories of certain characteristics (e.g., medium: graph, verbal, written) that Balcazar et al. (1985) did not. Nevertheless, both concluded that certain characteristics (e.g., private and public feedback, frequency of feedback) result in a change in performance and that these characteristics should be considered when applying feedback.

Although these reviews evaluated multiple characteristics of feedback, they did not address the type of feedback provided such as positive (e.g., "you are doing a great job") or negative (e.g., "your improvement is very slow") feedback, or the immediacy of feedback including immediate or delayed feedback. These two areas should be assessed to determine whether the immediacy and type of feedback provided contribute to a consistent change in behavior. For example, multiple studies have demonstrated that immediate feedback might be more effective than delayed feedback (Kulik & Kulik, 1988; Mason & Redmon, 1992; Scheeler, McKinnon, & Stout, 2012). These effects might be similar to reinforcement schedule effects (Prue & Fairbank, 1981). Additionally, some studies have found that positive feedback is more desirable than negative feedback and is likely to have an impact on an individual's behavior

(Jacobs, Jacobs, Feldman, & Cavior, 1973; Vallerand & Reid, 1988). This suggests that immediacy and type of feedback provided must also be considered in addition to the other characteristics described by the reviews mentioned above (Alvero et al., 2001; Balcazar et al., 1985). Therefore, the combination of immediate feedback provided publicly and privately on a daily basis might produce the greatest increases in performance.

Function of Feedback

Alvero et al. (2001) highlighted that a major goal of behavior analysis is to produce meaningful changes in behavior. In order to do that, one must understand the functional characteristics of feedback. Feedback is considered a consequence-based strategy and is likely to produce a change in behavior when it is introduced with antecedent strategies (Roane, Ringdahl, & Falcomata, 2015). However, research on feedback has been criticized for various reasons. There is controversy surrounding how feedback affects behavior and what specific characteristics make it effective. More specifically, some have argued that feedback functions in multiple ways and can be regarded as a discriminative stimulus, a conditioned reinforcer or punisher, and/or a motivating operation (Duncan & Bruwelheide,1985; Mangiapanello & Hemmes, 2015; Peterson, 1982).

The function of feedback can be discussed in relation to the GBG. The GBG has a feedback component that could be modified in certain ways depending on the population or setting. To recall, students are divided into two teams and the team who receives the lower number of points, wins the game. Before the game, the teacher or researcher states the rules of the game such as, sitting down nicely, and not speaking out of turn. During the game, the teacher or researcher places a hatch mark on a board (visual feedback) while announcing the team that is engaging in inappropriate behaviors (verbal feedback; e.g., “Team 1 gets a point for

not sitting down nicely"). The winning team receives a prize, such as lining up first for recess, or stickers. Due to the different functions that feedback can take, researchers (Duncan & Bruwelheide, 1986; Peterson, 1982) have suggested analyzing feedback procedures prior to implementing them in order to determine whether they function most effectively as a conditioned reinforcer or punisher (e.g., delivered immediately, by an experimenter, and privately), a discriminative stimulus (e.g., presented prior to a session), or motivating operation.

The feedback provided during the GBG might function as a conditioned reinforcer or punisher. That is, once students experience the GBG, the feedback provided during the game might begin to function as a reinforcer or punisher. For instance, when the teacher places a hatch mark on the board and provides verbal feedback to the team violating a rule (e.g., speaking out loud), this could serve as a conditioned punisher due to its history of being paired with losing the game and not being able to access a prize (e.g., stickers). Although most common GBG procedures involve hatch marks for breaking rules, one version, the Caught Being Good Game (Wright & McCurdy, 2011) provides hatch marks for following the rules. In this case, feedback might function as a conditioned reinforcer. In both cases, the feedback during the GBG provides students with information on their classroom performance which in turn, evokes a specific response.

On the other hand, the feedback during the GBG might function as a discriminative stimulus. During the GBG, the hatch marks and verbal feedback provided by the teacher is used in a punitive manner. For example, when hatch marks are presented for rule violations, this might indicate to students that they are less likely to receive reinforcement at the end of the game. However, when hatch marks are absent or few, this might indicate that reinforcement is more likely to be presented at the end of the game. Therefore, the feedback provided during the

game might function as an S-delta. That is, the hatch marks and verbal feedback provided might serve as an indicator of the availability of reinforcement. However, there are times that the feedback provided might not serve as a discriminative stimulus.

Feedback during the GBG might function as a motivating operation. When feedback is provided immediately after a specific response, the value of future consequences might change (Laraway, Snycerski, Michael, & Poling, 2003; Michael, 1982; Michael, 1993). A motivating operation is one in which an antecedent event alters the value of a reinforcer and increases the likelihood of behavior maintained by that reinforcer. In relation to the GBG, once a hatch mark is placed on the board, it might increase the punishing effects of additional hatch marks. The more hatch marks placed on the board, the less likely the students will contact reinforcement. Therefore, a hatch mark on the board increases the punishing value of additional hatch marks and abates any behaviors that produced those hatch marks in the past. As a motivating operation, the hatch marks and verbal feedback provided are presented contingent on the student's behaviors. It is less likely that these hatch marks are serving as discriminative stimuli in this case, because the addition of more hatch marks does not indicate that further punishment is available.

Generally speaking, it is difficult to determine what function feedback might play in producing changes in behavior. As such, it can be difficult to determine how to establish feedback as an effective behavior-change technique. Without a detailed analysis of the behavioral principles involved with feedback, it is difficult to ascertain how feedback might apply to the intervention or how to effectively analyze such effects. Normand, Bucklin, and Austin (1999) argued that the problem with feedback is that it includes an array of characteristics involving a variety of behavioral functions (e.g., reinforcers, discriminative stimulus, establishing operation) that are not often discussed or identified when applying an intervention.

According to DiGennaro et al. (2016), "a discussion of behavioral principles in *JOBM* articles could help readers maintain or refine their behavioral repertoire and aid in the adoption of function-based or indicated interventions" (p. 207). This statement applies to any application of feedback in that it should be consistently defined, even if it is not analyzed. Although there is little consensus on the function of feedback and how it changes behavior, there seems to be substantial evidence to support the use of certain characteristics of feedback towards improving performance. Therefore, applying characteristics of feedback associated with increases in performance to PA interventions might produce meaningful effects that are practical and useful (Alvero et al., 2001; Balcazar et al., 1985).

Feedback Characteristics and Physical Activity

There have been a number of PA studies that have used multiple characteristics of feedback (Donaldson & Normand, 2009; Hayes & Van Camp, 2015; Hustyi, Normand, & Larson, 2011; Kuhl et al., 2015). Most of these studies include feedback as a package intervention. For instance, Kuhl et al. (2015) analyzed the effects of two interdependent group contingencies (i.e., individual goal setting vs. group goal setting) based on the number of pedometer steps taken per day by third-grade children, while Hayes and Van Camp (2015) evaluated an intervention package that included feedback, reinforcement, self-monitoring, and goal setting, to increase PA exhibited by children during unstructured recess.

Both studies implemented various forms of feedback with many similarities. The similarities across the studies included the use of individual feedback (i.e., individual goal conditions compared to previous sessions without goals), unmasked pedometers to allow participants to view feedback on their own performance, rewards for meeting goals (e.g., crafts, small toys, extra recess), and verbal feedback which was provided immediately after specific

sessions in the form of praise for meeting a goal or encouragement to work on their goal for next time. Differences included group feedback (e.g., cumulative goal conditions) and public posting (e.g., visual boards), which Kuhl et al. (2015) applied during their study. Both studies (Hayes & Van Camp, 2015; Kuhl et al., 2015) applied characteristics of feedback that have been shown to produce consistent increases in performance (Alvero et al., 2001; Balcazar et al., 1985). That is, Kuhl et al. (2015) found that PA increased when they provided group and individual feedback, but the largest increase occurred in the individual feedback condition. Hayes and Van Camp (2015) found that the children took 47% more steps during intervention than in baseline. Both studies suggest that some form of feedback (e.g., public posting) might be useful towards increasing PA.

Public posting is a type of feedback that has decreased speeding on highways (Ragnarsson & Bjorgvinsson, 1991), inappropriate on-court behaviors by tennis players (Galvan & Ward, 1998), and inappropriate classroom behaviors such as running in hallways (Holland & McLaughlin, 1982). Furthermore, public posting has increased recycling behavior (Katzev & Mishima, 1992), bedmaking (Bacon-Prue, Blount, Hosey, & Drabman, 1980), academic performance (Van Houten, Hill, & Parsons, 1975), work performance (Nordstrom, Lorenzi, & Vance Hall, 1990), and sports performance (e.g., hockey, soccer, dancing; Anderson, Crowell, Doman, & Howard; Brobst & Ward, 2002; 1988; Quinn et al., 2017). Martin and Sharpe (2009) suggested that public posting could be used to increase motivation and participation during games, sports, and fitness activities. Because the use of public posting has produced consistent changes in behavior, using public posting to increase PA levels among individuals and groups might be beneficial. Incorporating public posting and other feedback characteristics (i.e., private

feedback, individual feedback, and group feedback provided immediately) during the GBG might be a promising step towards further increasing PA exhibited by children.

Feedback During the GBG

The GBG is an intervention package that has four core components: instruction, reinforcement, teams, and feedback. Instructions, in the form of providing classroom rules (e.g., stay in-seat, be on-task) are delivered by a teacher or researcher at the start of the GBG. During this time, the teacher or researcher describes the requirements to students on how to win the game and obtain a reward (Flower et al., 2014; Lynne et al., 2017). Additionally, positive reinforcement and differential reinforcement of low rate behavior (DRL) are implemented during the game. Positive reinforcement is used in the GBG in the form of providing rewards (e.g., stickers, extra recess) while DRL is provided to the group that engaged in the least amount of inappropriate behaviors (Litow & Pumroy, 1975). Furthermore, team distribution during the GBG might result in certain social contingencies arising. That is, if one team is losing, classmates might punish their teams' behavior and if the other team is winning, classmates might encourage their teams' behavior (Litow & Pumroy, 1975). Feedback during the GBG is typically provided immediately after the teacher observed a team member break one of the rules specified at the start of the game (Barrish et al., 1969; Donaldson et al., 2011; Pennington & McComas, 2017). In most cases, the teacher places a mark on the board (visual feedback) and states which team violated the rule (verbal feedback). The board is located where all students can see it throughout the game (Barrish et al., 1969; Groves & Austin, 2017; Salend, Reynolds, & Coyle, 1989). The announcement of the winning team and the reward provided at the end of the game might function as feedback. At times, the feedback in the GBG is modified to include an individualized component (Lutzker & White-Blackburn, 1979) or adjusted to incorporate

varied types of feedback (Lynne et al., 2017; Medland & Stachnik, 1972). For example, Lynne et al. (2017) used Class DoJo, an application that is used for tracking points and providing immediate feedback regarding students' behaviors through animation and sounds. Feedback is present during the GBG. Therefore, it is important to highlight, discuss, and evaluate its role within the GBG.

Most GBG studies have evaluated the effects of a complete intervention package compared to a baseline phase, but have not evaluated the effects that feedback has during the game (Barrish et al., 1969; Lynne et al., 2017; Swain, Allard, & Holborn, 1982), with only a few studies having conducted component analyses of the GBG (Foley, Dozier, & Lessor, 2018; Harris & Sherman, 1973; Medland & Stachnik, 1972). In fact, there have been only three component analyses conducted where each GBG component was analyzed, and the results are mixed (Foley et al., 2018; Harris & Sherman, 1973; Medland & Stachnik, 1972). Harris and Sherman (1973) found that the GBG with no feedback was just as effective as the GBG condition with feedback in reducing disruptive behavior in the classroom. Feedback included hatch marks (visual feedback) on a blackboard for disruptive behavior. This study concluded that feedback might not be a critical part of the GBG. However, Harris and Sherman (1973) found that students had experienced a long history of feedback before entering this condition, so carryover effects could have occurred from one condition to the next. Conversely, Medland and Stachnik (1972) found that feedback in the form of rules and lights (visual feedback) without the GBG decreased the frequency of problem behaviors in the classroom. This suggests that feedback might be an effective way to decrease problem behaviors and, hence, be an important part of the GBG. However, such conclusions are tentative because the experimenters only conducted the rules and lights phase over a short period (i.e., across nine sessions).

Still, those two studies (Harris & Sherman, 1973; Medland & Stachnik, 1972) suggest that the feedback component produced reductions in problem behaviors. And more recently, Foley et al. (2019) compared the effects of different GBG components (e.g., rules, feedback) before and after exposure to the GBG. They used a rule-plus-feedback condition in which feedback consisted of hatch marks as well as verbal statements to the team who violated a rule. Results showed that before the GBG was introduced, the individual components produced moderate reductions in disruptive behaviors, but there were larger decreases in disruptive behaviors after the students experienced the GBG (Foley et al., 2019). Overall, the literature reports inconsistent results in terms of the importance of feedback. Therefore, it might be important to determine the effectiveness of feedback during the GBG and whether the application of feedback with consistent characteristics (e.g., public feedback, individualized feedback) is likely to produce higher increases in appropriate behaviors such as PA.

Only one study has evaluated the effectiveness of feedback during the GBG (Wiskow et al., 2019). Wiskow et al. (2019) compared standard teacher contingencies (baseline), GBG, GBG without Feedback, GBG with visual feedback, GBG with vocal feedback, and GBG with visual and vocal feedback. Visual feedback included hatch marks on a whiteboard and vocal feedback included stating the rule that was violated contingent on the team engaging in disruptive behaviors. Results indicated that the GBG with vocal feedback and GBG with visual and vocal feedback conditions produced the largest decreases in disruptive behavior compared to the GBG without feedback condition. Wiskow et al. showed that feedback is an important component of the GBG. Therefore, using visual and vocal feedback similar to the standard GBG and evaluating the effects of feedback during the GBG on PA might add to the existing literature on the effectiveness of feedback during the GBG across behaviors.

Even though the rules of the GBG is largely the same whether targeting problem behavior or PA, the type of feedback and setting differs for studies that use the GBG to target PA (Galbraith & Normand, 2017; Jung et al., 2005). Jung et al. (2005) delivered feedback by publicly posting the accumulated points for each team. Conversely, Galbraith and Normand (2017) delivered verbal prompts (on a 3-min variable-time schedule) to remind teams to engage in PA, which could act as feedback. However, the prompt was not considered to be feedback as it did not specify anything about the teams performance, nor did it follow the target behavior of PA. Both studies announced the winning team at the end of the game, which might act as a form of feedback (Galbraith & Normand, 2017; Jung et al., 2005). Although both studies used some form of feedback, it is unclear whether this component had a notable effect on student's behaviors during the GBG. Additionally, it does not seem that all the characteristics of feedback that contribute to a consistent change in performance were applied across both studies (i.e., private feedback, public posting, group and individual feedback). Because no previous research has evaluated the effects of feedback during the GBG on PA, it would be beneficial to determine the usefulness of feedback during the game and without the game. This analysis might help clarify whether feedback is a significant component for the overall effectiveness of the GBG. That is, combining feedback similar to the standard GBG (Barrish et al., 1969) with the most beneficial aspects of feedback discussed during the GBG on PA could maximize the effectiveness of the GBG on PA.

Purpose of the Proposed Study

Few published studies have reported the effectiveness of feedback during the GBG (Foley et al., 2019; Harris & Sherman, 1973; Medland & Stachnik, 1972; Wiskow et al., 2019). Furthermore, no studies have compared the effects feedback has during the GBG on PA.

Evaluating feedback using characteristics (e.g., immediacy, delivery, mode of delivery) as described by Alvero et al. (2001) and Balcazar et al. (1985) might help determine whether feedback is a critical component of the GBG for increasing PA levels for children during recess. Visual and vocal feedback might produce larger increases in PA (Barrish et al., 1969; Wiskow et al., 2019). Galbraith and Normand (2017) only provided verbal prompts, with no visual or vocal feedback that is typically provided during the standard GBG (Barrish et al., 1969). Therefore, the purpose of the current study is to examine whether the addition of performance feedback enhances the effects of the GBG when used to increase the steps taken by elementary students playing on a school playground.

APPENDIX B: LITERATURE REVIEW REFERENCES

Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985-1998). *Journal of Organizational Behavior Management*, *21*(1), 3–29.

Anderson, D. C., Crowell, C. R., Doman, M., & Howard, G. S. (1988). Performance posting, goal setting, and activity-contingent praise as applied to a university hockey team. *Journal of Applied Psychology*, *73*(1), 87. doi:10.1037/0021-9010.73.1.87

Arundell, L., Ridgers, N.D., Veitch, J., Salmon, J., Hinkley, T., & Timperio, A. (2013). 5-year changes in afterschool physical activity and sedentary behavior. *American Journal of Preventive Medicine, 44,* 605–611. doi:10.1016/j.amepre.2013.01.029

Bacon-Prue, A., Blount, R., Hosey, C., & Drabman, R. S. (1980). The public posting of photographs as a reinforcer for bedmaking in an institutional setting. *Behavior Therapy*, *11*(3), 417–420. doi:10.1016/S0005-7894(80)80058-X

Balcazar, F., Hopkins, B. L., & Suarez, Y. (1985). A critical, objective review of performance feedback. *Journal of Organizational Behavior Management*, *7*(3–4), 65–89. doi:10.1300/J075v07n03_05

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*(2), 119–124. doi:10.1901/jaba.1969.2-119

Brobst, B., & Ward, P. (2002). Effects of public posting, goal setting, and oral feedback on the skills of female soccer players. *Journal of Applied Behavior Analysis, 35*(3), 247–257. doi:10.1901/jaba.2002.35-247

Brown, W. H., Pfeiffer, K. A., McIver, K. L., Dowda, M., Almeida, J. M., & Pate, R. R. (2006). Assessing preschool children's physical activity: The Observational System for Recording Physical Activity in children–preschool version. *Research Quarterly for Exercise and Sport*, *77*(2), 167–176. doi:10.5641/027013606X13080769704361

California Department of Education [CDE] (2013). 2013 Mathematics framework chapters. https://www.cde.ca.gov/ci/ma/cf/mathfwchapters.asp (accessed on December 3, 2017)

Cashwell, T., Skinner, C., & Smith, E. (2001). Increasing second-grade students' reports of peers' prosocial behaviors via direct instruction, group reinforcement, and progress feedback: A replication and extension. *Education and Treatment of Children, 24*(2), 161–175. doi:10.1002/j.2164-4683.1998.tb00405.x

Centers for Disease Control and Prevention [CDC] (2016). How much physical activity do children need. http://www.cdc.gov/physicalactivity/basics/children/ (accessed on August 1, 2017)

Centers for Disease Control and Prevention [CDC] (2017). Recess. https://www.cdc.gov/healthyschools/physicalactivity/recess.htm (accessed on September 30, 2017)

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson Education. ISBN-10: 0131421131

DiGennaro Reed, F. D., Henley, A. J., Rueb, S., Crabbs, B., & Giacalone, L. (2016). Discussion of behavioral principles in journal of organizational behavior management: an

update. *Journal of Organizational Behavior Management*, *36*(2–3), 202–209. doi:10.1080/01608061.2016.1200938

Donaldson, J. M., Fisher, A. B., & Kahng, S. (2017). Effects of the Good Behavior Game on individual student behavior. *Behavior Analysis: Research and Practice*, *17*(3), 207. doi:10.1037/bar0000016

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*. doi:10.1002/jaba.432

Donaldson, J. M., & Normand, M. P. (2009). Using goal setting, self-monitoring, and feedback to increase calorie expenditure in obese adults. *Behavioral Interventions*, *24*(2), 73–83. doi:10.1002/bin.277

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. doi:10.1901/jaba.2011.44-605

Duncan, P. K., & Bruwelheide, L. R. (1985). Feedback: Use and possible behavioral functions. *Journal of Organizational Behavior Management*, *7*(3–4), 91–114. doi:10.1300/J075v07n03_06

Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review*, *5*(4), 273–297.

Erwin, H. E., Ickes, M., Ahn, S., & Fedewa, A. (2014). Impact of recess interventions on children's physical activity—a meta-analysis. *American Journal of Health Promotion*, *28*(3), 159–167.

Escalante, Y., García-Hermoso, A., Backx, K., & Saavedra, J. M. (2014). Playground designs to increase physical activity levels during school recess: a systematic review. *Health Education & Behavior*, *41*(2), 138–144. doi:10.1177/1090198113490725

Fairclough, S. J., & Stratton, G. (2006). A review of physical activity levels during elementary school physical education. *Journal of Teaching in Physical Education*, *25*(2), 240–258. doi:10.1123/jtpe.25.2.240

Flower, A., McKenna, J.W., Bunuan, R.L., Muething, C.S., & Vega, R.V. (2014). Effects of the good behavior game on challenging behaviors in school settings. *Review of Educational Research, 84*(4)*,* 546–571. doi:10.3102/0034654314536781

Flower, A., McKenna, J., Muething, C. S., Bryant, D. P., & 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. doi:10.1177/0145445513507574

Foley, E., Dozier, C., & Lessor, A. (2019). Comparison of components of the Good Behavior Game in a preschool classroom. *Journal of Applied Behavior Analysis*, *52*(1), 84–104. doi:10.1002/jaba.506

Galbraith, L. A., & Normand, M. P. (2017). Step it up! using the good behavior game to increase physical activity with elementary school students at recess. *Journal of Applied Behavior Analysis*. doi:10.1002/jaba.402

Galvan, Z. J., & Ward, P. (1998). Effects of public posting on inappropriate on-court behaviors by collegiate tennis players. *The Sport Psychologist*, *12*(4), 419–426. doi:10.1123/tsp.12.4.419

Gortmaker, S. L., Peterson, K., Wiecha, J., Sobol, A. M., Dixit, S., Fox, M. K., & Laird, N. (1999). Reducing obesity via a school-based interdisciplinary intervention among youth: planet Health. *Archives of Pediatrics & Adolescent Medicine*, *153*(4), 409–418. doi:10.1001/archpedi.153.4.409

Greenwood, C. R., Hops, H., Delquadri, J., & Guild, J. (1974). Group contingencies for group consequences in classroom management: A further analysis. *Journal of Applied Behavior Analysis*, *7*(3), 413–425. doi:10.1901/jaba.1974.7-413

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. doi:10.1002/jaba.393

Hannon, J., & Brown, B. (2008). Increasing preschoolers' physical activity intensities: An activity-friendly preschool playground intervention. *Preventive Medicine, 46,* 532–536. doi:10.1016/j.ypmed.2008.01.006

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. doi:10.1901/jaba.1973.6-405

Harris, K. C., Kuramoto, L. K., Schulzer, M., & Retallack, J. E. (2009). Effect of school-based physical activity interventions on body mass index in children: a meta-analysis. *Canadian Medical Association Journal*, *180*(7), 719–726. doi:10.1503/cmaj.080966

Hayes, L. B., & Van Camp, C. M. (2015). Increasing physical activity of children during school recess. *Journal of Applied Behavior Analysis*, *48*(3), 690–695. doi: 10.1002/jaba.222

Heck, A., Collins, J., & Peterson, L. (2001). Decreasing children's risk taking on the playground. *Journal of Applied Behavior Analysis*, *34*(3), 349–352. doi:10.1901/jaba.2001.34-349

Hills, A. P., Dengel, D. R., & Lubans, D. R. (2015). Supporting public health priorities: recommendations for physical education and physical activity promotion in schools. *Progress in Cardiovascular Diseases*, *57*(4), 368–374. doi:10.1016/j.pcad.2014.09.010

Holland, E. L., & McLaughlin, T. F. (1982). Using public posting and group consequences to manage student behavior during supervision. *The Journal of Educational Research*, *76*(1), 29–34. doi:10.1080/00220671.1982.10885419

Houmanfar, R. (2013). Performance feedback: From component analysis to application. *Journal of Organizational Behavior Management, 33*(2), 85–88. doi:10.1080/01608061.2013.787002

Houten, R. V., Hill, S., & Parsons, M. (1975). An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. *Journal of Applied Behavior Analysis*, *8*(4), 449–457.

Houten, R., & Nau, P. A. (1983). Feedback interventions and driving speed: A parametric and comparative analysis. *Journal of Applied Behavior Analysis*, *16*(3), 253–281. doi: 10.1901/jaba.1983.16-253. doi:10.1901/jaba.1975.8-449

Howe, C. A., Freedson, P. S., Alhassan, S., Feldman, H. A., & Osganian, S. K. (2012). A recess intervention to promote moderate-to-vigorous physical activity. *Pediatric Obesity*, *7*(1), 82–88. doi:10.1111/j.2047-6310.2011.00007.x

Hustyi, K. M., Normand, M. P., & Larson, T. A. (2011). Behavioral assessment of physical activity in obese preschool children. *Journal of Applied Behavior Analysis*, *44*(3), 635–639. doi:10.1901/jaba.2011.44-635

Jacobs, M., Jacobs, A., Feldman, G., & Cavior, N. (1973). Feedback: II. The" credibility gap": Delivery of positive and negative and emotional and behavioral feedback in groups. *Journal of Consulting and Clinical Psychology*, *41*(2), 215. doi:10.1037/h0035100

Janssen, I., & LeBlanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 40.

Joslyn, P. R., Vollmer, T. R., & Hernández, V. (2014). Implementation of the good behavior game in classrooms for children with delinquent behavior. *Acta de Investigación Psicológica*, *4*(3), 1673–1682. doi:10.1016/S2007-4719(14)70973-1

Jung, J., Suroto, Fukugasako, Y., & Takahashi, T. (2005). The effects of group contingency strategy on students' behaviors in physical education classes in Korea using multiple baseline design. *International Journal of Sport and Health Science*, 3, 46–56. doi: 10.5432/ijshs.3.46

Katzev, R., & Mishima, H. R. (1992). The use of posted feedback to promote recycling. *Psychological Reports*, *71*(1), 259–264.

Kellam, S. G., & Anthony, J. C. (1998). Targeting early antecedents to prevent tobacco smoking: findings from an epidemiologically based randomized field trial. *American Journal of Public Health*, *88*(10), 1490–1495. doi:10.2105/AJPH.88.10.1490

Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., ... & Wilcox, H. C. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence*, *95*, S5–S28. doi:10.1016/j.drugalcdep.2008.01.004

Kelley, H., & Miltenberger, R. G. (2016). Using video feedback to improve horseback-riding skills. *Journal of Applied Behavior Analysis*, *49*(1), 138–147. doi: 10.1002/jaba.272

Kilanowski, C. K., Consalvi, A. R., & Epstein, L. H. (1999). Validation of an electronic pedometer for measurement of physical activity in children. *Pediatric Exercise Science*, *11*(1), 63–68. doi:10.1123/pes.11.1.63

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. doi:10.1177/0145445510392213

Kuhl, S., Rudrud, E. H., Witts, B. N., & Schulze, K. A. (2015). Classroom-based interdependent group contingencies increase children's physical activity. *Journal of Applied Behavior Analysis, 48*(3), 602–612. doi:10.1002/jaba.219

Kulik, J. A., & Kulik, C. L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, *58*(1), 79–97. doi:10.3102/00346543058001079

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*(3), 407–414. doi:10.1901/jaba.2003.36-407

Lee, A. M., Keh, N. C., & Magill, R. A. (1993). Instructional effects of teacher feedback in physical education. *Journal of Teaching in Physical Education*, *12*(3), 228–243. doi:10.1123/jtpe.12.3.228

Litow, L., & Pumroy, D.K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis, 8*(3), 341–347. doi:10.1901/jaba.1975.8-341

Lutzker, J.R., & White-Blackburn, G. (1979). The good productivity game: Increasing work performance in a rehabilitation setting. *Journal of Applied Behavior Analysis, 12,* 488. doi:10.1901/jaba.1979.12-488

Lynne, S., Radley, K. C., Dart, E. H., Tingstrom, D. H., Barry, C. T., & Lum, J. D. (2017). Use of a technology□enhanced version of the good behavior game in an elementary school setting. *Psychology in the Schools, 54*(9), 1049–1063. doi:10.1002/pits.22043

Mangiapanello, K. A., & Hemmes, N. S. (2015). An analysis of feedback from a behavior analytic perspective. *The Behavior Analyst*, *38*(1), 51–75. doi:10.1007/s40614-014-0026-x

Martin, M., & Sharpe, T. (2009). Using public posting as a motivation strategy in physical education, sport and adult exercise settings. *Strategies, 22*(3), 8–13. doi:10.1080/08924562.2009.10590818

Mason, M., & Redmon, W. K. (1992). Effects of immediate versus delayed feedback on error detection accuracy in a quality control simulation. *Journal of Organizational Behavior Management, 13,* 49–83. doi:10.1300/J075v13n01_05

McCurdy, B. L., Lannie, A. L., & Barnabas, E. (2009). Reducing disruptive behavior in an urban school cafeteria: An extension of the good behavior game. *Journal of School Psychology*, *47*(1), 39–54. doi:10.1016/j.jsp.2008.09.003

Mckee, D. P., Boreham, C. A., Murphy, M. H., & Nevill, A. M. (2005). Validation of the Digiwalker™ Pedometer for Measuring Physical Activity in Young Children. *Pediatric Exercise Science,17*(4), 345–352. doi:10.1123/pes.17.4.345

McKenzie, T. L., Sallis, J. F., Elder, J. P., Berry, C. C., Hoy, P. L., Nader, P. R., ... & Broyles, S. L. (1997). Physical activity levels and prompts in young children at recess: a two-year study of a bi-ethnic sample. *Research quarterly for exercise and sport*, *68*(3), 195–202. doi:10.1080/02701367.1997.10607998

McMurray,R.G., Harrell, J.S., Bangdiwala, S.I., Bradley, C.B., Deng, S., & Levine, A. (2002). A school-base intervention can reduce body fat and blood pressure in young adults. *Journal of Adolescent Health, 31,* 125–132. doi:10.1016/S1054- 139X(02)00348-8

Medland, M. B., & Stachnik, T. J. (1972). Good□behavior game: A replication and systematic analysis. *Journal of Applied Behavior Analysis*, *5*(1), 45–51. doi:10.1901/jaba.1972.5-45

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, *37*(1), 149–155. doi:10.1901/jeab.1982.37-149

Michael, J. (1993). Establishing operations. *The Behavior Analyst*, *16*(2), 191–206. doi:10.1007/BF03392623

Michael, J. (2000). Implications and refinements of the establishing operation concept. *Journal of Applied Behavior Analysis*, *33*(4), 401–410. doi:10.1901/jaba.2000.33-401

Miltenberger, R. G., Flessner, C., Gatheridge, B., Johnson, B., Satterlund, M., & Egemo, K. (2004). Evaluation of behavioral skills training to prevent gun play in children. *Journal of Applied Behavior Analysis*, *37*(4), 513–516. doi:10.1901/jaba.2004.37-513

Mota, J., Silva, P., Santos, M. P., Ribeiro, J. C., Oliveira, J., & Duarte, J. A. (2005). Physical activity and school recess time: differences between the sexes and the relationship between childrens playground physical activity and habitual physical activity. *Journal of Sports Sciences, 23*(3), 269–275. doi:10.1080/02640410410001730124

National Center for Education Statistics [NCES] (2005). State Education Reforms. https://nces.ed.gov/pubs2006/nutrition/tables/tab12.asp (accessed on September 30, 2017).

National Center for Education Statistics [NCES] (2015). State Education Reforms. https://nces.ed.gov/programs/statereform/tab5_14.asp (accessed on September 29, 2017).

National Physical Activity Plan Alliance. 2016 US report card on physical activity for children and youth. Columbia, SC; 2016.

Newby, T. J., & Robinson, P. W. (1983). Effects of grouped and individual feedback and reinforcement on retail employee performances. *Journal of Organizational Behavior Management*, *5*(2), 51–68.

Nordstrom, R., Lorenzi, P., & Hall, R. V. (1991). A review of public posting of performance feedback in work settings. *Journal of Organizational Behavior Management*, *11*(2), 101–124. doi:10.1300/J075v11n02_07

Normand, M., Bucklin, B., & Austin, J. (1999). The discussion of behavioral principles in JOBM. *Journal of Organizational Behavior Management*, *19*(3), 45–56.

Pate, R. R., Long, B. J., & Heath, G. (1994). Descriptive epidemiology of physical activity in adolescents. *Pediatric Exercise Science*, *6*(4), 434–447. doi:10.1123/pes.6.4.434

Pellegrini, A.D. (1995). *School recess and playground behavior: Educational and developmental roles.* Albany, NY: SUNY Press.

Pennington, B., & McComas, J. J. (2017). Effects of the good behavior game across classroom contexts. *Journal of Applied Behavior Analysis*, *50*(1), 176–180. doi:10.1002/jaba.357

Peterson, N. (1982). Feedback is not a new principle of behavior. *The Behavior Analyst*, *5*(1), 101–102. doi:10.1007/BF03393144

Prue, D. M., & Fairbank, J. A. (1981). Performance feedback in organizational behavior management: A review. *Journal of Organizational Behavior Management*, *3*(1), 1–16. doi:10.1300/J075v03n01_01

Quinn, M., Miltenberger, R., Abreu, A., & Narozanick, T. (2017). An intervention featuring public posting and graphical feedback to enhance the performance of competitive dancers. *Behavior Analysis in Practice*, *10*(1), 1–11. doi:10.1007/s40617-016-0164-6

Ragnarsson, R. S., & Björgvinsson, T. (1991). Effects of public posting on driving speed in Icelandic traffic. *Journal of Applied Behavior Analysis*, *24*(1), 53–58. doi:10.1901/jaba.1991.24-53

Ramstetter, C. L., Murray, R., & Garner, A. S. (2010). The crucial role of recess in schools. *Journal of School Health*, *80*(11), 517–526. doi:10.1111/j.1746-1561.2010.00537.x

Reid, D. H., & Parsons, M. B. (1996). A comparison of staff acceptability of immediate versus delayed verbal feedback in staff training. *Journal of Organizational Behavior Management*, *16*(2), 35–47. doi:10.1300/J075v16n02_03

Reilly, J. J., & Kelly, J. (2011). Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: Systematic review. *International Journal of Obesity*, 35, 891–898. doi:10.1038/ijo.2010.222.

Reiner, M., Niermann, C., Jekauc, D., & Woll, A. (2013). Long-term health benefits of physical activity–a systematic review of longitudinal studies. *BMC Public Health*, *13*(1), 813. doi:10.1186/1471-2458-13-813

Roane, H. S., Ringdahl, J. E., & Falcomata, T. S. (Eds.). (2015). *Clinical and organizational applications of applied behavior analysis*. Academic Press.

Salend, S., Reynolds, C., & Coyle, E. (1989). Individualizing the good behavior game across type and frequency of behavior with emotionally disturbed adolescents. *Behavior Modification, 13*, 108–126. doi:10.1177/01454455890131007

Sallis, J. F., Patterson, T. L., McKenzie, T. L., & Nader, P. R. (1988). Family variables and physical activity in preschool children. *Journal of Developmental & Behavioral Pediatrics*, *9*(2), 57–61. doi:10.1097/00004703-198804000-00001

Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. *Medicine & Science in Sports & Exercise*, *32*(5), 963–975. doi:10.1097/00005768-200005000-00014

Scheeler, M. C., McKinnon, K., & Stout, J. (2012). Effects of immediate feedback delivered via webcam and bug-in-ear technology on preservice teacher performance. *Teacher Education and Special Education*, *35*(1), 77–90. doi:10.1177/0888406411401919

Simons-Morton, B. G., Parcel, G. S., Baranowski, T., Forthofer, R., & O'Hara, N. M. (1991). Promoting physical activity and a healthful diet among children: results of a school-based intervention study. *American Journal of Public Health*, *81*(8), 986–991. doi:10.2105/AJPH.81.8.986

Stokes, J. V., Luiselli, J. K., Reed, D. D., & Fleming, R. K. (2010). Behavioral coaching to improve offensive line pass-blocking skills of high school football athletes. *Journal of Applied Behavior Analysis*, *43*(3), 463–472. doi:10.1901/jaba.2010.43-463

Stone, E. J., McKenzie, T. L., Welk, G. J., & Booth, M. L. (1998). Effects of physical activity interventions in youth: review and synthesis. *American Journal of Preventive Medicine*, *15*(4), 298–315. doi:10.1016/S0749-3797(98)00082-8

Swain, J. J., Allard, G. B., & Holborn, S. W. (1982). The good toothbrushing game: A school-based dental hygiene program for increasing the toothbrushing effectiveness of children. *Journal of Applied Behavior Analysis*, *15*, 171–176. doi:10.1901/jaba.1982.15-171

Sy, J. R., Gratz, O., & Donaldson, J. M. (2016). The good behavior game with students in alternative educational environments: interactions between reinforcement criteria and scoring accuracy. *Journal of Behavioral Education*, *25*(4), 455–477. doi:10.1007/s10864-016-9257-0

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. doi: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. doi:10.1177/0145445503261165

Tudor-Locke, C., Craig, C. L., Rowe, D. A., Spence, J. C., Tanaka, S., Blair, S. N., ... & Hatano, Y. (2011). How many steps/day are enough? for children and adolescents. *International Journal of Behavioral Nutrition and Physical Activity, 8*(78), 1–14. doi:10.1186/1479-5868-8-78

Vallerand, R. J., & Reid, G. (1988). On the relative effects of positive and negative verbal feedback on males' and females' intrinsic motivation. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement*, *20*(3), 239. doi:10.1037/h0079930

Van Houten, R., Hill, S., & Parsons, M. (1975). An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. *Journal of Applied Behavior Analysis*, *8*(4), 449–457. doi:10.1901/jaba.1975.8-449

Van Sluijs, E. M., McMinn, A. M., & Griffin, S. J. (2007). Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *BMJ*, *335*(7622), 703. doi:10.1136/bmj.39320.843947.BE

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. doi:10.1177/0145445516644220

Warburton, D. E., & Bredin, S. S. (2017). Health benefits of physical activity: a systematic review of current systematic reviews. *Current Opinion in Cardiology*, *32*(5), 541–556. doi:10.1097/HCO.0000000000000437

Wiskow, K., M., Matter, A., & Donaldson, J. M. (2019). The Good Behavior Game in preschool classrooms: An evaluation of feedback. *Journal of Applied Behavior Analysis, 52*(1), 105–115. doi:10.1002/jaba.500

World Health Organization [WHO] (2017). Physical Activity. http://www.who.int/mediacentre/factsheets/fs385/en/ (accessed on August 1, 2017)

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. doi:10.1177/1098300711421008

Zerger, H. M., Miller, B. G., Valbuena, D., & Miltenberger, R. G. (2017). Effects of student pairing and public review on physical activity during school recess. *Journal of Applied Behavior Analysis*, *50*(3), 529–537. doi:10.1002/jaba.389

APPENDIX C: LIST OF POTENTIAL REINFORCERS

1. First choice to line up for recess, lunch, library, etc.
2. Have an extra 5 minutes on the computer.
3. Have an extra 5 minutes of recess.
4. Have lunch in the classroom with the teacher.
5. First choice to pick an activity during activity time.
6. Dismiss 2 minutes early from school.
7. Be given a “raffle ticket” that the students can write their name on and throw it into a bowl for prize drawings.
8. Gets to pick the activity for the entire class.
9. Get to sit with your best friend for the rest of the day.
10. Be given a “skip one night of homework” pass.
11. Get 1 extra credit point on an assignment.
12. Receiving a prize from a treasure box (e.g., stickers, pencils, sharpeners).

APPENDIX D: MATH ASSESSMENT

1. Jasmine scored the following number of points in 5 dart games. What is the mean (average) of these numbers?

   88, 96, 112, 135, 144

   a. 56 b. 88 c. 112 d. 115

2. Alex is on the Eagles Bowling Team. His scores for the last 12 games are shown below. What is the mean (average) of these numbers?

   90, 103, 110, 95, 105, 110,
   90, 112, 110, 96, 94, 110

   a. 90 b. 102 c. 104 d. 110

3. The chart below shows the scores for five students on last week's math test. The test was out of 10 points. Find the mean (average) score.

Meagan	Kelly	Julia	Ingrid	Matt
9	8	9	8	8

   a. 9 b. 8 c. 7 d. 6

4. The classroom mean (average) for the last English test was 84%. Jenna scored an 80% on the test. Did Jenna score:
   a. Above average
   b. Below average
   c. At average
   d. None of the above

5. The group mean steps (average) during recess was 1300. Steve took 1500 steps during recess. Did Steve perform:
   a. Above average
   b. Below average
   c. At average
   d. None of the above

6. The group mean steps (average) during recess was 1800. Mike took 1200 steps during recess. Did Mike perform:
    a. Above average
    b. Below average
    c. At average
    d. None of the above

7. The group mean steps (average) during recess was 1500. Jenna took 1500 steps during recess. Did Jenna perform:
    a. Above average
    b. Below average
    c. At average
    d. None of the above

APPENDIX E: SOCIAL VALIDITY QUESTIONNAIRE

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
1. The Step it Up! Game by itself increased physical activity.	1	2	3	4	5
2. The combination of feedback and Step it Up! Game increased physical activity.	1	2	3	4	5
3. The intervention seemed simple to implement.	1	2	3	4	5
4. The study took too long.	1	2	3	4	5
5. I enjoyed the overall experience.	1	2	3	4	5

6. Which part of the intervention did you like the most (e.g., Step it Up! Game, Step it Up! Game and feedback)?

7. Was there any portion of the experience you would like to change? If yes, what change(s) do you recommend?

8. We welcome any additional comments/suggestions. Thank you.

Note. Some parts are adapted from Galbraith and Normand (2017)