

Western New England University

Digital Commons @ Western New England University

Doctoral Dissertations - College of Arts and Sciences

College of Arts and Sciences

2023

Evaluation of a competition using a fitness tracker for increasing physical activity

Chelsea Headquist

Western New England University

Follow this and additional works at: <https://digitalcommons.law.wne.edu/casdissertations>

Recommended Citation

Headquist, Chelsea, "Evaluation of a competition using a fitness tracker for increasing physical activity" (2023). *Doctoral Dissertations - College of Arts and Sciences*. 22.

<https://digitalcommons.law.wne.edu/casdissertations/22>

This Dissertation is brought to you for free and open access by the College of Arts and Sciences at Digital Commons @ Western New England University. It has been accepted for inclusion in Doctoral Dissertations - College of Arts and Sciences by an authorized administrator of Digital Commons @ Western New England University.

**EVALUATION OF A COMPETITION USING A FITNESS TRACKER FOR
INCREASING PHYSICAL ACTIVITY**

By

Chelsea B. Hedquist

M.S., Western New England University, 2013

Dissertation Defense

Submitted to the Department of Psychology

and the College of Arts and Sciences at

Western New England University in partial

fulfillment of the requirements for the

Degree of Doctor of Philosophy

Dr. Eileen M. Roscoe, Committee Chair

Dr. Jason C. Bourret, Committee Member

Dr. Rachel H. Thompson, Committee Member

Dr. Jon W. Pinkston, Committee Member

Date: January 31st, 2023

Abstract

Physical activity has been associated with numerous health benefits, including improved sleep and decreased risk of diabetes. However, individuals rarely meet the Centers for Disease Prevention and Control recommendation of 150 min per week of moderate-to-vigorous-physical activity (MVPA). Therefore, identifying new strategies for increasing MVPA is an important goal for promoting health and wellness. One helpful tool for measuring physical activity is a fitness tracker that automatically records heart rate, step count, exercise minutes, calories, and stand hours. In addition to these features, some fitness trackers offer a competition feature that allows users to challenge other individuals to compete for points gained through physical activity. The purpose of the current study was to evaluate the effects of the competition feature on MVPA among typically developing adults. If the competition alone was insufficient, additional components, including differential reinforcement of alternative behavior, differential observing response, and a contingency contract, were evaluated. Participants ($n = 20$) were paired with another adult ($n = 10$ pairs) and competitions were arranged. In experiment 1, the competition increased levels of MVPA to criteria performance for 5 of 6 participants. For the remaining participants ($n = 14$), additional intervention components (i.e., differential observing response and contingency contracts) were subsequently combined with the competition feature to evaluate MVPA in experiment 2.

Keywords: physical activity, fitness tracker, competition, automated feedback

Evaluation of a Competition using a Fitness Tracker for Increasing Physical Activity

Physical activity is an important component to living a healthy lifestyle. Engaging in moderate-to-vigorous-physical activity (MVPA) is associated with lowered risk for Type 2 diabetes cardiovascular disease, obesity, anxiety, and other chronic diseases (CDC, 2014). Other benefits include improvements to overall health, sleep, and mood (CDC, 2014). Examples of MVPA include taking a brisk walk, swimming, jumping rope, and playing soccer. The Centers for Disease Control and Prevention (CDC) recommends 60 min of MVPA per day for children and 150 min per week of MVPA in combination with strength training for adults (CDC, 2015). The World Health Organization (WHO) estimates that 60% of adults do not meet this recommendation and that over 5 million premature deaths worldwide could be prevented if individuals engaged in the recommended amount of physical activity (WHO, 2020). Although small increases in low-intensity physical activity have health benefits, more extended durations of MVPA (i.e., those that increase heart rate engagement have greater health benefits (AHA, 2018). Therefore, identifying ways to measure and increase MVPA is important.

One way to increase physical activity is to arrange a competition between participants with the goal of engaging in the most physical activity. Normand and Burji (2020) arranged a competition between students in a third-grade classroom to compete for the most steps using the Step it UP! game. Across baseline and intervention conditions, the participants wore a pedometer on their wrist, which recorded their total step count. During baseline, there were no programmed consequences in effect. During the competition condition, participants were divided into two teams based on their baseline levels of responding (e.g., individuals with high step counts were paired with individuals with low step counts during baseline). Prior to the start of the

competition, the experimenter reviewed the rules, “At the end of PE, the team that took the most steps will win the game and get a prize.” When the PE class was over, step count was calculated for each participant and team, then the winner was announced. Step count increased during the competition compared to baseline for 14 of 18 total participants. Participants and PE staff both indicated that they enjoyed the game and would play it again in the future. Although these results are encouraging, it is unclear whether an increase in step count corresponded with increases in MVPA. Future research should evaluate a competition using other types of fitness trackers that allow for other health related measures to be taken.

A convenient way to measure MVPA during a competition is by using an automated fitness tracker (e.g., Fitbit, Apple Watch, Whoop, Halo, etc.) that continuously records the individual’s heart rate (Apple, 2022). The tracker (e.g., Fitbit, Apple Watch) is generally worn on the wrist or a finger (i.e., Oura) and provides a continuous data summary of the wearer’s physical activity within and across days. For example, the Apple Watch continuously displays three different colored rings that correspond with three different physical activity goals. The green ring displays minutes of MVPA, the blue ring displays the number of hours of at least 1 min of standing, and the red ring displays the percentage of the individual’s self-identified movement goal (based on calories burned from physical activity engagement) that was met. In summary, fitness trackers offer a convenient method for measuring physical activity progress and other health-related behavior.

Some fitness trackers offer a competition, or a function that allows two people to challenge one another to engage in more health-related behavior to win the most points (Apple, 2021). For example, the Apple Watch offers seven-day competitions that can be arranged between two people. Competitors compete to obtain the most points during a week, and the

winner receives a badge that is displayed on their personal Fitness Application. A point is earned each time the user meets the day's exercise goal, the stand goal, or the movement goal, which is displayed on the fitness tracker as filling up or closing the full circle for the goal. These rings and the associated response measures can be viewed from the watch or from the Fitness Application on the user's iPhone. Taken together, the competition and measurement system the fitness tracker offers, might increase users' MVPA to the CDC health recommendations. A comprehensive literature search returned no previous studies that have evaluated the utility of the competition feature of the Apple Watch. Therefore, the effects of this feature on MVPA remains unknown.

There may be conditions under which the competition alone is insufficient to increase MVPA to the CDC recommendations, in these cases, additional intervention components could be added to enhance treatment outcomes. For example, differential reinforcement of alternative behavior (DRA), differential observing response (DOR), or a contingency contract (CC) could be added to the competition. DRA involves the delivery of a reinforcer for an alternative behavior and extinction for another response. An example of DRA was illustrated in a study by Hayes and Van Camp (2015). The authors evaluated a multiple-component intervention that included wearing a fitness tracker for increasing step count and MVPA of third graders during recess. During baseline, the experimenters covered the Fitbit screens with tape to prevent participants from viewing their step count data. During the intervention, the experimenters removed the tape covering the Fitbit screen so that participants could view their step count data. In addition, each day the experimenters instructed the participants to view their Fitbit step count during each 20-min session. At the end of the 20-min session, participants received praise for meeting goals or feedback and encouragement for not meeting goals. The automated feedback from the Fitbit plus

DRA (differential praise) and feedback from the experimenter resulted in increases in step count for all participants. For five of the six participants, step count and average MVPA increased during intervention relative to baseline. DRA may be an effective treatment intervention to add to a competition when the competition alone is ineffective.

A second example of a component that could be added to the competition is a DOR. For example, responding toward a sample stimulus before a stimulus array is presented to increase the probability of attending to specific features of the sample stimulus (Grow & LeBlanc, 2013). Establishing a contingency for a DOR involves providing a reinforcer to increase a target response that may facilitate attending to the stimuli that could function as discriminative stimuli. For example, during the competition, there are points, leaderboard, and other competition related stimuli that may signal the availability of reinforcement. However, these stimuli can only function as discriminative stimuli if the participants attend to them. For example, Lillie and Tiger taught college undergraduates to emit a DOR while learning print-to-braille relations. The DOR involved copying a sample stimulus (i.e., the correct number and location of braille dots) then matching the sample to the corresponding stimulus in an array. Participants acquired the braille-to-print match-to-sample responses after the DOR training. The DOR contingency may have established the stimulus as a discriminative stimulus. Therefore, inclusion of a DOR contingency may facilitate the competition related stimuli (i.e., exercise min and points) functioning as a discriminative stimulus.

A third treatment component that may help facilitate increases in physical activity engagement during a competition is CC. A CC is a contingency (contract) made between two or more people that agree about a target response and corresponding reinforcer (Cooper et al., 2020). CC (e.g., specific exercise-related goal agreed upon by the participant and experimenter)

has been used to increase various forms of physical activity engagement including aerobic activities (Wysocki et al., 1979), step count (Washington et al., 2014), and step count virtually (Valbuena et al., 2015).

A type of CC that has been effective to increase step count is a deposit contract (Donlin Washington et al., 2016). In a deposit contract, participants deposit money or belongings into an account monitored by the experimenter prior to intervention. During the intervention, participants can earn money or belongings back by achieving predetermined goals. For example, Donlin Washington et al. (2016) evaluated the effects of a deposit contract on step count. During baseline, participants were instructed to wear Fitbits and report their daily step count totals to experimenters (via phone call text, or email), and no programmed contingencies were in effect. Prior to the intervention, the experimenter developed a step count goal based on a percentile schedule from participants' baseline performance. Participants were randomly placed in a deposit ($n = 10$) or no deposit group ($n = 9$). The deposit group was required to deposit \$25 to be earned back, in addition, experimenters matched the \$25 meaning participants could earn up to \$50 for meeting goals. The participants in the no deposit group ($n = 10$) did not have to deposit money. Rather, they could earn up to \$50 for meeting goals. During both interventions, the experimenter informed participants of the contingency (i.e., the specific step count required for receiving their money) and participants could earn small sums of money for meeting daily goals. There were increases in step count for both groups. The experimenters reported that the deposit contract was advantageous to keep costs down for experimenters relative to the no deposit group. One limitation was that MVPA was not measured or reported. Although step count corresponds with MVPA, it is not a direct measure of active min, and the CDC guideline is based upon active min (i.e., duration) rather than step count (i.e., frequency). When considering the competition feature

of the Apple Watch, it may be possible for someone to win a competition, but not meet the CDC guidelines (e.g., levels of physical activity are low but greater than the opponent's). In these cases, adding contingency contracts or deposit contracts may result in an effect treatment package for increasing MVPA.

In summary, if the competition alone is ineffective to increase physical activity engagement to meet the CDC recommendations, then additional treatment components could be added to enhance treatment outcomes. DRA (Hayes & Van Camp, 2015), DOR (Lillie & Tiger, 2017), and CC (Donlin Washington et al., 2016) have been shown to be effective to increase target behavior and may be effective to increase physical activity engagement. DRA could be added for the winner of the competition, DOR could be added to increase the saliency of the competition components, or CC could be added as an additional contingency for meeting the CDC recommendations.

Competitions may be an advantageous treatment option as they can be implemented remotely (competitors do not need to be in the same location or go to a lab or gym), data can be shared easily, and there is flexibility with the physical activity type, duration, and time of day. Given the flexibility with which a competition can be implemented, this could be a meaningful approach to increasing physical activity and approaching the CDC recommendations (i.e., MVPA). The purpose of the present analysis was to evaluate the utility of a competition using a fitness tracker on MVPA in 20 individuals. In Study 1, the effects of the competition alone were evaluated. In Study 2, the effects of the competition alone and with additional treatment components when the competition alone was ineffective, were evaluated.

General Method

Participants and Materials

Participants ($n = 20$) were typically developing adults. To recruit participants, the experimenter sent an email to a group of teachers with information about physical activity and the importance of engaging in the duration (150 min) of MVPA per week recommended for adults by the CDC. Additional details in the email were that many adults do not meet the CDC recommended duration of MVPA, the risks for not engaging in physical activity, and the benefits of engaging in physical activity. At the end of the email, the experimenter asked if (the recipient) was interested in participating in a research project to increase their physical activity and if they currently owned an apple watch. Individuals who expressed interest in participating were asked for their consent to share their Apple Health data and Fitness Application data with the experimenters for up to six months. All participants were required to supply their own Apple Watch, cell phone, and computer. After the consent form was completed, participants were asked to complete an initial indirect assessment (IA) via Forms (Microsoft Office 365). This IA was completed by all participants ($M = 11$ min). Once completed, the experimenter reviewed the IA results and then scheduled the first meeting via Zoom (Banyani, 1995). During this meeting, the experimenter reviewed the consent form with the participant, provided a brief project overview, and asked the participants if they had any questions about the IA they completed or the project. Next, participants were asked to add the experimenter to their contact list and their Fitness Application, which allowed for data sharing with the Apple Watch. The experimenter explained that their personal activity data would be hidden from view by the participant, whereas the participant's daily physical activity summary data would be available for view by the experimenter. Last, participants were asked to share their historical data from the Health Application and to send the file to the experimenter. These data included exercise min, movement calories, and stand hours. This condition was conducted to assess performance prior

to the start of the first meeting and rule out potential reactivity effects of the experimenter viewing participant data.

Response Measurement

The primary dependent variable was exercise, defined as moderate-to-vigorous physical activity (MVPA) based on the heart rate output by the Apple Watch. Exercise min were summarized as duration in min, and data were collected by the Apple Watch (Apple, 2022). Additional dependent variables included movement calories, stand hours, and points. All dependent variables were extracted from the Fitness Application, viewable from the Apple Watch or an iPhone. Movement calories corresponded to the calories expended from movement. Stand hours indicated the number of hours that the wearer stood for at least 1 min.

An additional dependent variable was points. Points were automatically calculated during competitions (i.e., the person with the most points in a competition wins). During baseline conditions, the experimenter calculated points by using the same formula that the Apple Watch automatically uses. The Formula is based on the three previously noted dependent variables (i.e., exercise min, movement, and stand) and the proportion of each goal that was met. Exercise and stand goals were the same for all participants (see Table 1). By contrast, the movement calorie goal differed across participants because it was individualized and determined prior to participation in the study. The specific formula used to calculate points was:

$$\left(\frac{\text{Active Calories Output}}{\text{Active Calories Goal}} + \frac{\text{Exercise Output}}{\text{Exercise Goal}} + \frac{\text{Stand Output}}{\text{Stand Goal}} \right) * 100.$$

For instance, if a participant's movement goal (active calorie goal) was 600 a day and the participant burned 300 calories in a day, the participant would have achieved 50% of their goal (*300 active calories [output] divided by 600 active calories [goal] multiplied by 100*) and been awarded 50 points from their active calorie goal for that day. Alternatively, if the participant's

movement goal was 400 calories a day and the participant burned 500 active calories in a day, then the participant would have achieved *125% of the goal (500 active calories [output] divided by 400 active calories [goal] multiplied by 100)* and been awarded 125 points from the calorie goal for that day. Percentages were rounded to the nearest tenth. This same formula was also applied to calculate points for activities related to movement and stand goals each day. Once points were calculated for each goal, they were added together for a total score and rounded to the nearest whole number. Apple Watch has a cap of 600 points per day; thus, one cannot obtain more than this cap by engaging in the target behavior more often. Experimenters reviewed and graphed points and exercise min throughout the study. Because points were only calculated and displayed by the Apple Watch and Fitness Application during a competition, participants could only view their points during intervention phases.

The Fitness Application provides data on many health measures and displays progress toward an individual's goals across time. Goals were individually determined prior to the start of the study (e.g., when the user set up their Apple Watch). To ensure that a participant's goal remained consistent during their participation in the Study, the experimenter asked participants not to change their goals for the duration of the study.

Reliability

For data collection and reliability purposes, the experimenters downloaded and saved the participants' daily summary data from the Fitness Application. Trained graduate students, currently enrolled in a master's program in applied behavior analysis, entered the daily summary data into an excel document. The experimenters compared those records with their own separate excel file. If a discrepancy was observed, the experimenter asked the observer to review the summary data to ensure they were looking at the correct day and review if there was a mistake.

This occurred on two occasions and was an error in data entry (e.g., the exercise min were entered at 61, but the correct response was 16). Agreement was scored in this way for all three studies. For Study 1, IOA was completed for 100% of sessions for participants 2 and 3, 73% of sessions for participants 1 and 4, and 49% of sessions for participants 6 and 7. For Study 2, IOA was calculated for 42% of sessions for participants 1 and 4, 44% for participants 5 and 8, 57% of sessions for participants 2 and 3, and 34% of sessions for participants 6 and 7. For Study 3, IOA was calculated for 100% of sessions for participants 1 and 3, 98% for participants 4 and 5, and 80% for participants 2 and 6. Mean IOA was 100% across all participants for Study 1, Study 2, and Study 3.

Procedures

IAs

Participants were asked to complete IAs at the onset of the study, after each treatment condition, and at the end of the study. The experimenter sent all IAs to participants electronically via Microsoft Forms and asked participants to complete at their earliest convenience. Once completed, the experimenter received an email notification from Microsoft Forms that an IA had been submitted. The initial IA was conducted prior to the start of the study and had three purposes. The first purpose was to ensure that the individual could safely participate in the study by completing a questionnaire. The Physical Activity Readiness Questionnaire (PAR-Q; Thomas et al., 1992) was sent via email. The PAR-Q included medical questions that could make it unsafe for them to participate. For example, two of the questions were “In the past month, have you had chest pain when you were not doing physical activity? Do you know of any other reason why you should not do physical activity?” If the participant responded affirmatively to any of the PAR-Q questions, indicating a medical reason not to participate, they were excluded from

participation in the study (this occurred for one individual). The second purpose was to gain information about participants' level of physical activity and use of Apple Watch. The third purpose was to assess social validity and satisfaction with their current levels of physical activity.

Initial Meeting

The initial meeting was conducted remotely (via Zoom) after the initial IA was completed and was 10 to 15 min in duration. During this meeting, the experimenters reviewed details in the consent form, how to send Apple Health Data, and how to add someone to the Fitness Application to share their data. At the end of the meeting, the experimenter instructed participants to, "Please continue to engage in activity as your normally would, I will reach out again soon."

Pre-condition Review Meetings

Prior to initiating a new treatment condition, the experimenter met with the participant for a brief (2 to 5 min) meeting via Zoom. During these meetings, the experimenter noted the antecedents and consequences that the participant would experience in the upcoming intervention condition. The experimenter also asked the participant to share data with their competition opponent through the Fitness Application and initiate a competition.

Apple Watch

The participant wore his or her own Apple Watch across baseline and intervention conditions. The Apple Watch sends automatic notifications in the form of textual prompts that appear on the Apple Watch screen. The content of these messages varied based on participant responding. For example, if the participant had not moved in a while, the watch may have prompted, "Keep moving, you are almost to your goal!" If the participant has not stood for close

to an hour, the watch may have prompted, “Time to stand. Stand up and move a little for one minute.” These prompts were part of the Apple Watch interface and could not be modified.

Participants were asked to wear their watch across all waking hours, except when it was being charged.

Design

A reversal design was used to assess the effectiveness of intervention components. For an intervention to be deemed effective, individualized performance criteria were used. These criteria were determined from participants’ mean duration of exercise min (MVPA) during the baseline condition (described below). If the participant engaged in less than 30 exercise min per day, then the following criteria were used: (a) a mean of 22 exercise min per day (to meet CDC recommendations) and (b) 10% or greater increase in exercise min compared to the previous baseline. For participants who engaged in more than 30 min of exercise per day during baseline, criteria included: (a) a mean of at least 30 exercise min per day, and (b) 5% or greater increase in exercise min compared to the previous baseline. Because this latter group of participants exhibited baseline performance that met the CDC recommendations for physical activity, a less stringent criterion for percentage of change in exercise min was used for them. Although these individuals exhibited high baseline levels, they were included as participants because they expressed interest in increasing their levels of physical activity.

Pre-baseline data

Participants were asked to share their health data via email prior to the start of the study. Experimenters reviewed 14 nonconsecutive days of data days to permit exclusion of days in which participants did not wear their watches. These data were subsequently organized and exported into Excel using a programming application that was developed in Python for this

purpose. The health data (e.g., exercise min, stand hr, active calories) were reviewed prior to sending the participants consent forms. The purpose of this phase was to have a measure of participants' physical activity prior to formal participation in the study, as this would allow an assessment of potential reactivity effects (e.g., increased or decreased physical activity) associated with experimenters reviewing participant's data each day.

Baseline

This condition began immediately after the initial meeting (described above). The experimenter stated, "Please continue to engage in activity as your normally would. I will reach out again soon." There were no programmed consequences in effect. The experimenter reviewed participants' data during this condition to form participant pairs who showed similar baseline performance (i.e., exercise min) for the competition condition.

Four of six participants in Study 1 were paired based on obtaining a similar mean of points during baseline. The participant who earned the fewest points was ranked in position one, the participant who earned the next fewest points was ranked in position two, and so on. Next, participants were ranked according to their position, such that those in positions one and two were matched, and those in positions three and four were matched. Because the Apple Watch calculated points only during competitions, the experimenter calculated points across conditions to ensure a consistent unit of measure. The remaining two participants in Study 1 and all participants in Study 2 and Study 3 were paired based on mean exercise min during baseline. Participants were ranked by the number of exercise min they obtained from the lowest number of minutes to the highest number of minutes. Participants in adjacent positions were then paired (position one with two, three with four, and so on). After the experimenter identified participant

pairs based on their baseline performance, these pairs remained consistent throughout all post-baseline conditions.

Study 1

Competition

Prior to the start of this condition, the experimenter held a pre-condition review meeting with each participant. During this meeting, the experimenter instructed the participant to add their opponent using the opponent's phone number (provided by the experimenter) to their contacts as "Opponent 1." The identity of the participants' opponent was not disclosed, and participants had no direct interaction with their opponent. Next, participants were asked if the added opponent was previously a contact. The experimenter instructed the participants to not discuss the study with their opponent at any time. After adding their opponent, the experimenter informed the participant that they would momentarily receive an invitation to share data from their opponent via the Fitness Application and that they should accept the request. By sharing data, participant pairs could view one another's activity summary data (i.e., progress towards move, exercise, and stand goals, number of steps, distance from steps, and logged workouts).

During this meeting, the experimenter described what the competition entailed. Participants who had previously experienced a competition, could opt out of receiving these instructions. The experimenter explained that there would be a seven-day competition in which points could be earned for meeting their goals of moving, exercising, and standing (depicted on the watch as the closing of their activity rings). Participants were informed that they could view their opponent's progress each day (e.g., move, exercise, stand goals, and total points). The experimenter explained that a maximum of 600 points could be earned each day and a total of 4,200 points per week. In addition, the experimenter noted that there was a leaderboard that

showed which opponent was winning and that they could expect to receive frequent automatic prompts on their watch regarding their performance in relationship to their opponents. (e.g., “You are only behind by a few points, keep going!”) The experimenter reviewed how to earn points, how points were calculated, and that the winner would be awarded a badge through the application.

At the end of the meeting, the experimenter prompted participant pairs to initiate a competition. The first participant in the pair was prompted to initiate a competition by opening the Fitness Application and selecting “opponent 1”, followed by “compete with opponent 1.” The second participant in the pair was instructed to accept the request when received, and the competition automatically commenced the subsequent day (e.g., if the second participant accepted the competition invite on a Monday, then the competition began on Tuesday). The experimenter then stated, “Please try your best to beat your opponent. Do you have any questions?” No further instructions or programmed consequences provided. The day after the competition ended, the experimenter contacted the participant and asked the participant to remove their opponent from their fitness application. Once an opponent was removed, participants could no longer access their opponent’s health data.

Results

The results for the first competition pair (participants 1 and 2) are depicted in Figure 1. Both participants exhibited low levels of exercise min during pre-baseline ($M = 5.3$ min and $M = 10.7$ min per day for participants 1 and 2, respectively). In the baseline condition, exercise min decreased from pre-baseline levels for both participants ($M = 1.3$ min and $M = 7.4$ min a day for participants 1 and 2, respectively). During the initial competition condition, exercise min increased for both participants to performance criteria ($M = 52.1$ min a day, 3907.7% increase

from baseline for participant 1; $M = 27.4$ min a day, 269.2% increase from baseline for participant 2). There was a decrease in responding for both participants ($M = 3.4$ min and $M = 16.5$ min a day for participants 1 and 2, respectively) during the second baseline condition. In the second competition condition, exercise min increased to performance criteria levels for both participants ($M = 45.6$ min a day, 1244.7% increase from baseline for participant 1; $M = 23.4$ min per day, 53.8% increase from baseline for participant 2). In summary, the competition alone was effective in increasing exercise min to performance criteria for participants 1 and 2.

The results for competition pair two (participants 3 and 4) are depicted in Figure 2. During the pre-baseline condition, exercise min were greater than 30 min per day for both participants ($M = 45.9$ exercise min and $M = 67.6$ exercise min per day for participants 3 and 4, respectively). There was a decrease in exercise min from the pre-baseline to baseline for participant 3 ($M = 38.6$ min per day) and an increase for participant 4 ($M = 78.6$ min per day). Because the only difference between pre-baseline and baseline conditions was sharing physical activity data and having it reviewed by the experimenters, the increase in exercise min during this initial baseline from may have been due to potential reactivity. Because these participants consistently engaged in 30 exercise min per day in the first baseline condition, their performance criteria were more stringent: (a) a mean of 30 exercise min per day, and (b) 5% or greater increase in exercise min compared to the previous baseline.

During the first competition condition, participant 3 met performance criteria ($M = 91.9$ min per day, 138.1% increase from baseline), but participant 4 did not ($M = 87.6$ min per day, 4.2% increase from baseline). During the second baseline condition, exercise min decreased for both participants ($M = 65.6$ and $M = 57.5$ min per day for participants 3 and 4, respectively). During the second competition condition, both participants met performance criteria ($M = 98.7$

exercise min per day, 52.2% increase from baseline for participant 1; $M = 99.3$ exercise min per day, 11.4% increase from baseline for participant 4). Given the less than 5% increase from baseline for participant 4 during the first competition, an additional baseline and competition were completed to assess whether similar levels of performance observed in the second competition condition would occur. During the following baseline condition, exercise min decreased ($M = 47.4$ exercise min and $M = 58.7$ exercise min per day for participants 3 and 4, respectively). During the third competition condition, both participants met performance criteria ($M = 100.4$ exercise min, 52.8% increase from baseline for participant 3; $M = 102$ exercise min, 42.5% increase from baseline for participant 4). The competition was effective to increase exercise min to performance criteria for both competition pair one (participants 3 and 4).

The results for competition pair three (participants 5 and 6) are depicted in Figure 3. Although these participants knew one another prior to the start of the study, both agreed not to discuss the competition or the study throughout the duration of involvement. However, it is possible that the intervention was effective given the features that are associated with knowing your competitor. During pre-baseline, exercise min were high for both participants ($M = 99.9$ exercise min and $M = 91.3$ exercise min per day for participants 5 and 6, respectively). During baseline, there was a decrease in responding from the pre-baseline for both participants ($M = 97.8$ and $M = 88.6$ exercise min per day for participants 5 and 6, respectively). Because these participants consistently engaged in greater than 30 exercise min per day in the first baseline condition, their performance criteria were more stringent (a) a mean of 30 exercise min per day, and (b) 5% or greater increase in exercise min compared to the previous baseline.

During the first competition, there was a decrease in responding for participant 5 ($M = 90.1$ exercise min, -7.78% change) and an increase to performance criteria for participant 8 ($M =$

10.4 exercise min per day, 15.6% change). During the second baseline condition, there was a decrease in exercise min for both participants ($M = 85.3$ exercise min per day and $M = 84.9$ exercise min per day for participants 5 and 6, respectively). During the second competition condition, exercise min increased, and the performance criteria were met for both participants ($M = 106.6$ exercise min, 24.9% increase from baseline for participant 5; $M = 103.4$ exercise min, 21.8% increase from baseline for participant 6). Given the high levels of exercise min during the competition far exceeding the CDC recommendations (e.g., 150 min a week or about 22 min a day), no further replications were completed. These results were considered effective for participant 6, but not for participant 5 (only met criteria during second replication).

At the end of the analysis, all participants completed a final IA and provided feedback on their level of physical activity and satisfaction. When asked, “how satisfied are you with your Apple Watch helping you meet your physical activity goals when you are competing vs not competing?” participants responded similarly. Three participants responded that they were satisfied with how their watch facilitated meeting goals. One participant noted that they were satisfied, but did not think the watch made a difference outside of the competition (when they did find it helpful to increase their physical activity). Two participants noted that they were more motivated during competitions compared to time when they were not competing. When asked if they would engage in competitions in the future, three participants responded yes and three participants responded sure (or potentially). Future research should continue to evaluate the parameters associated with the competition that would result in participants being more interested in competing after the study is over.

Study 2

Results of Study 1 indicated that the competition alone was effective for 5 of the 6 participants. However, there may be conditions in which the competition alone is ineffective and additional treatment components may be necessary to add. When the competition alone is ineffective to increase exercise min to performance criteria, several treatment options may increase the effectiveness of the competition (e.g., DRA, DOR, and DOR plus DC).

Participants and Materials

Fourteen individuals who did not participate in Study 1, were included in Study 2.

Procedures

Competition

This condition was identical to the competition condition described for Study 1, including the pre-condition review meeting.

Competition Plus DRA

During the pre-condition review meeting, the same steps described for the competition alone condition were reviewed again. In addition, the experimenter noted that the winner of the competition (i.e., the participant who received the highest number of points at the end of the competition) would earn a monetary prize (\$30 via Venmo). At the end of the meeting the experimenter stated, "Please try your best to beat your opponent. There will now be a \$30 prize for the winner of the competition. Do you have any questions?" No questions were asked, although some participants commented that they would like to win the money. The following day, the competition began, and participants could view one another's data again. At the end of the condition, the experimenter reached out to the winner to send the \$30.

Competition Plus DOR

If the competition did not increase exercise min to performance criteria levels, then the DOR was added. During the pre-condition review meeting, the experimenter restated the rules of the competition. In addition, the DOR requirement was explained to the participant. The DOR required the participant to take a photograph (from their watch or their phone) of the competition data for the day and send it to the experimenters each evening by 9:00 pm. In addition, they were required to send the experimenter a text message each evening that included a summary of the total points and exercise minutes they and their opponent earned for that day. For example, the text would include the following details, “me: 35 min, 400 pts; opponent: 28 min, 378 pts”. If participants did not send the experimenter a text message with photograph and data summary report by 9:00 pm, the experimenters sent them a prompt via text message to do so (e.g., “Just a friendly reminder to please send your daily screen shot and text”). The experimenter also instructed participants to check that their watch’s home screen was configured to continuously display the three data rings. At the end of the meeting, the experimenters stated, “Please try your best to beat your opponent. Make sure to send your data each day and take a screen shot of the competition to send it to the experimenters. Any questions?”

Competition Plus DOR Plus Contingency Contract

If the competition plus DOR condition was ineffective to meet performance criteria, then a deposit contract was evaluated. The rules of the competition were the same as previously described. Prior to beginning this condition, participants were sent a description of the condition and procedures via email.

At the pre-condition review meeting, the experimenter described the procedures to the participant, noting that they could earn their money back for meeting their daily goals, and that they would receive their earned money at the end of the competition (Donlin Washington et al.,

2016; Jarvis & Dallery, 2017; Stedman-Falls & Dallery, 2020). To meet CDC recommendations, the contingency requirement was that participants earned a mean of 22 min of exercise per day. For example, if a participant engaged in 60 exercise min on day 1, the goal for day 1 would be met. If they subsequently engaged in 10 min on day 2, the goal for day 2 would be met because the mean across the two days would be 35 min. The mean calculation would be adjusted for include exercise min from previous days and the current day of the competition. Using an adjusting mean for the contingency requirement gave participants greater flexibility in achieving their goals (i.e., they did not need to engage in exactly 22 exercise min each day). After participants sent their text with their scanned and typed daily data summary, the therapist responded via text by stating, “Thank you. You did (or did not) meet your goal of a mean of 22 exercise min a day for today. You are now earning \$XX.” The text was accompanied by a short video clip (a gif) that displayed a seven-piece pie chart. Each piece of the pie corresponded to the day of the competition and had a corresponding dollar amount on it, illustrating the total amount deposited and the total amount that was earned or lost. For example, if the participant met their goal the first day, the pie would contain all 7 pieces (and show \$70). If the participant did not meet their goal on the second day, then the pie would depict one of the seven pieces being removed and \$60. If participants did not meet exercise goals, the deposited money would be donated to a charity of the experimenters choosing. The experimenter returned participants earn money to them via Venmo the day after the competition ended.

Results

Figure 4 depicts the results for participants 1 and 2. During the pre-baseline condition, both participants exhibited low levels of exercise min ($M = 6.9$ min and $M = 12.5$ min per day for participants 1 and 2, respectively). In the initial baseline condition, exercise min increased from

pre-baseline levels for participant 1 ($M = 22.8$ min per day) but remained low for participant 2 ($M = 10.3$ min per day). The increase observed for participant 1 may have been due to potential reactivity associated with the experimenter reviewing the participant's data. During the initial competition condition, exercise min increased to performance criteria for both participants ($M = 55.4$ min per day, 142.7% increase from baseline for participant 1; $M = 42$ min per day, 351.1% increase from baseline for participant 2). During the following baseline condition, exercise min decreased for both participants ($M = 31.5$ and $M = 9.5$ min per day for participants 1 and 2, respectively). During the second competition condition, neither participant met the performance criteria. Exercise min decreased for participant 1 ($M = 16.5$ min per day, 37.8% decrease from baseline) and increased for participant 2 ($M = 20.6$ min per day, 114.7% increase from baseline). Although participants met the percentage change criterion, they did not meet the criterion of an average of 22 exercise min a day. Given that these results were inconsistent with the first competition phase, a third competition condition was conducted to evaluate the effects of the competition alone. In the following baseline condition, exercise min remained low for participant 1 ($M = 13.9$ min per day) and decreased for participant 2 ($M = 8.1$ min per day). When the competition was implemented a third time, exercise minutes remained low for both participants and neither participant met performance criteria ($M = 11$ min per day, 21.1% decrease from baseline for participant 1; $M = 10.1$ min per day, 24.9% increase from baseline for participant 2). Because the competition was not effective in increasing exercise min to performance criteria for both participants, the addition of a DRA component was subsequently assessed. During the initial competition plus DRA condition, there was a decrease in exercise min for participant 1 ($M = 5.2$ min per day, 67.1% decrease from baseline) and an increase for participant 2 ($M = 49.6$ min per day, 364% increase from baseline). Participant 1 did not meet performance criteria,

whereas Participant 2 did meet criteria. During the subsequent baseline condition, exercise min maintained at low levels for participant 1 ($M = 9.3$ min per day) and decreased for participant 2 ($M = 8.0$ min per day). In the second competition plus DRA condition, participant 1 did not meet performance criteria ($M = 4.3$ min per day, 38.8% decrease from baseline), whereas participant 2 did meet performance criteria ($M = 36.4$ min per day, 355.4% increase from baseline). In summary, the competition alone was not consistently effective to increase exercise min to criteria levels for the second competition pair. The addition of DRA to the competition was effective in increasing exercise min to performance criteria for participant 2. However, this was not effective in increasing exercise min for participant 1.

Figure 5 depicts the results for participants 3 and 4. During the pre-baseline condition, exercise min was low for both participants ($M = 12.1$ and $M = 9.6$ min per day for participants 3 and 4, respectively). During baseline, exercise min maintained at low levels ($M = 12.3$ and $M = 8.9$ min per day for participants 3 and 4, respectively). During the initial competition condition, exercise min increased for both participants ($M = 18.9$ min per day, 53.9% increase from baseline for participant 3; $M = 27.1$ min per day, 203.7% increase from baseline for participant 4). However, only participant 6 met performance criteria. When the competition was no longer in effect during the subsequent baseline condition, exercise min decreased for both participants ($M = 9.6$ min and $M = 15.1$ min per day, for participants 3 and 4, respectively). When the competition condition was conducted a second time, neither participant met the performance criteria. Participant 3 showed increases in exercise min relative to baseline ($M = 16.7$ min per day, 71.4% increase from baseline) but not to criterion levels, and participant 4 showed a decrease exercise min ($M = 6.1$ min per day, 59.3% decrease from baseline). During the following baseline condition, there was an increase in exercise min for both participants ($M =$

20.1 min and $M = 7.1$ min per day for participants 3 and 4, respectively). Both participants did not meet the criterion of an average of 22 exercise min a day but did meet the percentage change criterion. Given the inconsistent outcome for participant 4 a third competition condition was implemented to further assess its effects. During the third competition condition, participant 3 met performance criteria ($M = 28.3$ min per day, 41.6% increase from baseline), whereas participant 4 did not ($M = 7.6$ min per day, 11.7% increase from baseline). During the subsequent baseline condition, exercise min increased for participant 3 ($M = 35.5$ min per day) and remained stable for participant 4 ($M = 7.3$ min per day). Given that competition alone was not consistently effective for both participants, DRA was subsequently added to the competition. During the competition plus DRA condition, participant 3 met performance criteria ($M = 40.0$ min per day, 13.5% increase from baseline) and participant 4 did not ($M = 6.5$ min per day, -10.5% decrease from baseline). During the return to baseline, exercise min decreased for both participants ($M = 23.0$ and 5.27 min per day, for participants 3 and participant 4, respectively). In the second competition plus DRA condition, exercise min increased to performance criteria for participant 3 ($M = 30.7$ min per day, 33.5% increase from baseline); but not for participant 4 ($M = 6.9$ min per day, 30.1% increase from baseline). The addition of DRA to the competition was effective to increase exercise min to performance criteria for participant 3. However, this was not effective to increase exercise min for participant 4.

Figure 6 depicts the results for participants 5 and 6. During pre-baseline, both participants engaged in low levels of exercise min ($M = 8.6$ exercise min and $M = 12.3$ exercise min per day for participants 5 and 6, respectively). In baseline, there were low levels of exercise min for both participants ($M = 7.3$ exercise min and $M = 10.3$ exercise min per day for participants 5 and 6, respectively). During the first competition condition, exercise increased for both participants (M

= 14.3 exercise min, 94.8% increase from baseline for participant 5; $M = 22.1$ min, 114.3% increase from baseline for participant 6). This increase was sufficient for only participant 5 to meet performance criteria. During the second baseline condition, there was a decrease in exercise for both participants ($M = 7.8$ min and $M = 15.4$ min for participants 5 and 6, respectively). During the second implementation of the competition, neither participant met the performance criteria. For participant 5, there was an increase in exercise min, but not to sufficient levels ($M = 15.3$ min, 97.2% increase from baseline). For participant 6, exercise decreased ($M = 11.7$ exercise min, -24.1% change). Because the competition alone was not effective for participants 5 and 6, the competition combined with a DOR was subsequently evaluated. During the third baseline condition, both participants showed decreases in exercise ($M = 6.3$ min and $M = 9.9$ min for participants 5 and 6, respectively). During the competition plus DOR, there was an increase in exercise min for both participants, but not to performance criteria ($M = 16.0$ min, 153.4% increase from baseline for participant 5; $M = 19.7$ min, 98.7% increase from baseline for participant 6). During a return to baseline, there was a decrease in responding for both participants ($M = 7.2$ and 7.1 min for participants 5 and 6, respectively). During the second implementation of the competition plus DOR, there was an increase in exercise min for both participants, but not to performance criteria ($M = 13.9$ min, 48.0% increase from baseline for participant 5; $M = 10.1$ min, 30.9% increase from baseline for participant 6). Although there was an increase in exercise min during the competition and competition plus DOR conditions, exercise min did not meet performance criteria under either intervention condition.

Results for participants 7 and 8 are depicted in Figure 7. During pre-baseline, participant 7 did not meet the CDC recommendations ($M = 21.9$ min per day), whereas participant 8 did meet the CDC recommendations ($M = 43.0$ min per day). For both participants, there was a

decrease in responding from naturalistic baseline to baseline ($M = 18.8$ and 38.5 min per day for participants 7 and 8, respectively). Because participant 8 engaged in a mean of over 30 exercise min during the initial baseline, the percentage of change criterion was a 5% increase from baseline. For participant 7, who engaged in less than 30 exercise min during baseline, the percentage of change criterion was a 10% increase from baseline. During the initial implementation of the competition, exercise increased to performance criteria for both participants ($M = 63.9$ min per day, 240.6% increase from baseline for participant 7; $M = 70.6$ min per day, 83.3% increase from baseline for participant 8). During the return to baseline, there was a decrease in exercise min for both participants ($M = 22.5$ and $M = 29.8$ min per day for participants 7 and 8, respectively). During the second implementation of the competition, there was a decrease in exercise min for participant 7 ($M = 19.6$ min per day, -12.7% decrease from baseline) and an increase in exercise to performance criteria for participant 8 ($M = 37.28$ min per day, 25.33% increase from baseline). Participant 7 disclosed that they were traveling and forgot their watch charger. When the watch ran out of battery on day 5, this participant ended the competition manually on their phone. Given this potential confound, a third condition of competition alone was evaluated after a return to baseline ($M = 18$ min 26.5 min per day for participants 7 and 8, respectively). During the third competition condition, there was an increase to performance criteria for participant 7 ($M = 34.29$ min per day, 90.48% increase from baseline), but a decrease in responding for participant 8 ($M = 22.83$ min per day, -16.88% decrease from baseline). During a return to baseline, exercise decreased for both participants ($M = 27.0$ min and 22.0 min per day for participants 7 and 8, respectively). Because the competition alone did not result in sufficient increases in exercise min across two replications, the effects of competition plus DOR was subsequently assessed. During the competition plus DOR, there was

an increase in exercise min to performance criteria for both participants ($M = 53$ min per day, 49% increase from baseline for participant 7; $M = 32.7$ min per day, 30.2% increase from baseline for participant 8). Due to circumstances unrelated to the study, one of the participants had to drop out, so this pair did not experience a replication of competition plus DOR.

Results for participants 9 and 10 are depicted in Figure 8. During the pre-baseline condition, there were low levels of exercise min for both participants ($M = 4.6$ min and $M = 13.9$ min per day for participants 9 and 10, respectively). When the initial baseline condition was implemented, exercise min remained low ($M = 4.1$ min and $M = 10.3$ min per day for participants 9 and 10, respectively). During the competition plus DOR condition, exercise min increased for both participants ($M = 11.4$ min per day, 179.9% increase from baseline for participant 9; $M = 12.3$ min per day, 18.9% increase from baseline for participant 10). Although exercise min increased during this condition, the increase was not to the criteria performance. Therefore, the addition of a contingency contract component was subsequently evaluated. During the second baseline condition, there was a decrease in exercise min for participant 9 ($M = 8.7$ min per day), and an increase for participant 10 ($M = 19.5$ min per day). During the first implementation of the competition plus DOR with contingency contract, exercise min increased, and performance criteria were met for both participants ($M = 22.9$ min per day, 161.7% increase from baseline for participant 9; $M = 27.1$ min per day, 28.0% increase from baseline for participant 10). Both participants received their \$70 back for meeting their goals. During the third baseline condition, exercise min decreased for both participants ($M = 5.1$ min and $M = 14$ min per day for participants 9 and 10, respectively). During the second implementation of the competition plus DOR with contingency contract, both participants showed an increase in exercise min ($M = 17.6$ min per day, 251.4% increase from baseline for participant 9; $M = 23.3$

min per day, 39.9% increase from baseline for participant 10). However, only participant 3 met the performance criteria and earned their \$70 back. Participant 9 earned \$50 back. In summary, the competition plus DOR with contingency contract was effective for participant 10 only.

Figure 9 depicts the results for participants 11 and 12. During pre-baseline, both participants engaged in moderate and low levels of exercise min ($M = 24.6$ min and $M = 13.9$ min per day for participants 11 and 12, respectively). During baseline, exercise decreased for participant 11 ($M = 15.6$ min per day) and maintained at low levels for participant 12 ($M = 14.1$ min per day). During the first intervention phase, competition plus DOR, there was an increase in exercise min for both participants. However, only participant 11 met performance criteria ($M = 26.6$ min per day, 69.9% increase from baseline), participant 12 did not ($M = 17.1$ min per day, 17.5% increase from baseline). Given this outcome, this condition was conducted a second time for only participant 11; participant 12 experienced the addition of the contingency contract. During the return to baseline, exercise increased for participant 11 ($M = 34.6$ min per day), and exercise decreased for participant 12 ($M = 15.6$ min per day). During the second competition plus DOR condition, exercise min remained high but did not meet performance criteria ($M = 36.1$ exercise min, 4.2% change). Because participant 11 was less than 1% away from meeting the performance criteria, and there was increase from baseline to intervention, participant 11 continued to experience the competition plus DOR. Further, participant 11 let the experimenter know that they started a new exercise routine from pre-baseline to baseline. During the first competition plus DOR with contingency contract condition, exercise min increased and met the performance criteria for participant 12 ($M = 33.6$ min per day, 114.6% increase from baseline). During the return to baseline, there was an increase in exercise min for participant 11 ($M = 41.6$ min per day) and a decrease in exercise min for participant 12 ($M = 21.2$ min per day). During

the third competition plus DOR condition, exercise decreased for participant 11 ($M = 41.4$ min per day, -0.29% decrease from baseline). During the second competition plus DOR with the contingency contract condition, there was an increase in exercise min and participant 12 met performance criteria ($M = 24.7$ min per day, 17.7% increase from baseline). In summary, participant 11 had an overall increase in exercise min across the second baseline condition and competition plus DOR conditions with a daily average of over 30 exercise min per day. Given the new exercise routine and overall increase in exercise min, the results of the intervention are unclear. For participant 12, the competition plus DOR alone was ineffective, and the addition of the contingency contract was effective in increasing exercise to performance criteria.

Figure 10 shows the results for participants 13 and 14. During pre-baseline, there were low levels of exercise min for both participants ($M = 13.3$ per day min and $M = 9.2$ min per day for participants 13 and 14, respectively). When baseline was implemented, exercise decreased for participant 13 ($M = 10.8$ min per day) and increased for participant 14 ($M = 13.3$ min per day). During the first competition plus DOR condition, exercise decreased for participant 13 ($M = 9.6$ min per day, -11.7% decrease from baseline) and increased and met the performance criteria for participant 14 ($M = 26.6$ min per day, 176.0% increase from baseline). Given these outcomes, the contingency contract was added to the subsequent competition plus DOR condition for participant 13 only. During the return to baseline, exercise decreased for both participants ($M = 4.3$ min and $M = 20.8$ min for participants 13 and 14, respectively). During the first competition plus DOR with contingency contract condition, exercise min increased to performance criteria for participant 13 ($M = 34.0$ min per day, 674.7% increase from baseline). During the second competition plus DOR condition, exercise min remained low and performance criteria were not met for participant 14 ($M = 20.4$ min per day, -1.6% decrease from baseline).

Therefore, the addition of the contingency contract was evaluated for participant 14. During the return to baseline, there was a decrease in exercise for both participants ($M = 5.1$ min per day and $M = 15.4$ min per day for participants 13 and 14, respectively). During the following competition plus DOR with contingency contract condition, exercise increased and met the performance criteria for both participants ($M = 30.4$ min per day, 496.7% increase from baseline for participant 13; $M = 26.7$ min per day, 73.6% increase from baseline for participant 14). During the return to baseline, exercise decreased for both participants ($M = 3.8$ min per day and $M = 16.8$ min per day for participants 13 and 14, respectively). During the subsequent competition plus DOR with contingency contract condition, exercise min increased but not to the performance criteria for participant 13 ($M = 19.7$ min per day, 418.4% increase from baseline), whereas exercise increased and met the performance criteria for participant 14 ($M = 25.6$ min per day, 52.7% increase from baseline). Although the competition plus DOR with contingency contract increased exercise min for both participants, only participant 2 consistently met the performance criteria.

When the study completed, only 13 of the 14 participants completed the final IA. The remaining participant declined to respond when prompted a second time. When participants were asked, “how satisfied are you with your Apple Watch helping you meet your physical activity goals when you are competing vs not competing?” 10 participants responded that they were satisfied or very satisfied. The remaining three participants responded that they exercised more during the competition. When asked if they would engage in competitions in the future, 12 participants responded yes. One participant noted that they wanted to buy their spouse an Apple Watch just to set up competitions with him. Two participants responded no that they would not engage in future competitions. The participants that responded no also noted that they prefer to

work out on their own schedule and not have the interference of a competition, the other participant noted they did not think they would be able to find anyone to complete with. Future research could consider arranging participants to compete (with the absence of the experimenter facilitating) after the study is over, if participants are still interested in competing.

General Discussion

The efficacy of the competition feature of the Apple Watch was evaluated in the present analysis ($n = 20$). During Study 1, the competition alone was effective to increase exercise min to performance criteria for five of six participants. During Study 2, when DRA was added (i.e., the winner of the competition received \$30) for four participants, two additional participants met the performance criteria. The final two participants that experienced DRA did not meet performance criteria. When DOR was added to the competition for four participants, there was an increase in exercise min, but not to performance criteria. For the final three groups, the competition plus DOR plus the deposit contract was effective to increase performance to criteria for three of five participants.

The results replicated and extended Krentz et al. (2016), who found that DRA (token delivery for each lap walked, tokens exchanged at the end of the session) was effective in increasing walking laps for four of the five participants. The DRA arrangement used in the current study differed from that used by Krentz et al. Whereas Krentz et al. offered a reinforcer to each participant following each session based on number of laps walked, the DRA contingency in the current study involved delivery of a monetary reward (i.e., \$30) at the end of the 7-day competition contingent on winning the competition. This DRA contingency was effective in increasing exercise min to performance criteria for only two of the four participants who experienced the condition. The DRA contingency may have been more effective if the

reward was delivered more frequently (e.g., \$30 each day to the participant who performed the best that day) or contingent on a specific performance criterion being met (e.g., having at least 22 min of exercise per day) by each participant. It is also possible that the use of a larger monetary prize at the end of the study (e.g., \$500) would have been more effective. Future research could evaluate these modifications to facilitate increases in MVPA to meet the CDC recommendations.

Another explanation for why DRA was not as effective as Krentz et al. might be due to the target behavior used in the current study. The criteria used in the current study was a mean of 22 exercise min a day and a 10% increase in exercise min from baseline (for participants who engaged in less than 30 exercise min a day in baseline) or a mean of 30 exercise min a day and a 5% increase in exercise min from baseline (for participants who engaged in more than 30 exercise min a day during baseline). By contrast, Krentz et al. did not measure MVPA or use performance criteria to assess the efficacy of the DRA contingency; the efficacy of their intervention was based on laps walked. Therefore, our criteria for deeming an intervention as effective may have been more stringent than that used by Krentz et al. Despite participants having flexibility with exercise activity and timing, some participants did not meet the performance criteria. Future research should continue to evaluate different performance criteria, perhaps based on baseline levels of responding and individual goals.

The evaluation of DOR replicated and extended previous DOR research by Lillie and Tiger (2017), who demonstrated the utility of a DOR (i.e., matching braille character patterns) to increase skill acquisition of braille characters. For all participants, the DOR was successful in increasing acquisition of braille characters to mastery. In the current analysis (Study 2), the DOR required the participant to type and send an electronic message to the experimenter at the end of each day of the competition. In the message, the participant was required to list the number of

exercise min and points they and their opponent earned that day and to attach a scanned image of these data from the leaderboard. Although the addition of the DOR increased exercise min for all participants, the increases were not sufficient for meeting performance criteria for two participants. Future research could consider modifying the DOR to further increase the saliency of the relevant stimuli and enhance the effects of the competition. For example, participants could be required to send multiple electronic messages that include current data summaries from the leaderboard at different times across the day to increase attending to the stimuli more frequently and earlier in the day. Increasing attending to the data earlier in the day may facilitate the data functioning as a discriminative stimulus when there is a greater opportunity to engage in the target response (i.e., MVPA).

When the competition alone and competition plus DOR were insufficient to increase exercise min to performance criteria, an additional deposit contract was evaluated in Study 3. We extended research by Donlin Washington et al. (2016), who demonstrated the efficacy of a deposit contract for step count (participants in the deposit contract group met 70.9% of daily step count goals), by using a deposit contract to increase MVPA to performance criteria levels for three of five participants. In both studies, the money was delivered at the end of the intervention phase. However, in the current study we requested a larger sum of money to be deposited (i.e., \$70) compared to a smaller sum of money (i.e., \$25) in the Donlin Washington et al. analysis. Another difference is that the \$25 was matched by experimenters for a total of \$50 that could be earned over time. In the present study, we did not match participant deposits, but this is a worthwhile area of future research. Another future direction to increase efficacy of the deposit contract, could be to investigate different reinforcer delivery schedules (e.g., daily instead of at

the end of the intervention phase) and different deposit amounts (e.g., smaller than \$25 or larger than \$70).

Our study extended Dolin Washington et al. by targeting average daily MVPA rather than daily step count. In Donlin Washington et al., not meeting the daily step count goal resulted in loss of money for the day. By contrast, in Study 3 of the current paper, we extended the application of the deposit contract by setting the goal to engage in a mean duration of 22 exercise min a day across a 7-day period. Using a mean calculation for the target behavior (in this case, exercise min) offered participants more flexibility in how they achieved the criterion for the deposit contract. For example, if the participant engaged in 22 min of exercise on day 1, they would have met the goal for that day ($M = 22$ exercise min). If they engaged in 0 min on day 2, they would not have met criterion for that day ($M = 11$ exercise min a day). If they engaged in 44 exercise min on day 3, they would have met the criterion for that day ($M = 22$ exercise min per day). Participants could continue to earn money each day by engaging in different durations of exercise min across days that were most convenient, which may have facilitated meeting performance criterion for the 7-day competition. Participants 1 and 2 who did not meet performance criteria demonstrated an increase in exercise min; they had a large percentage change between baseline and intervention, suggesting that the deposit contract had utility in increasing their exercise min. Future research should compare different types of goals (e.g., daily v. average daily target responses), to determine the relative efficacy of various approaches to deposit contracts.

In addition to comparing different types of goals, future research should evaluate the mechanisms responsible for behavior changes during deposit contracts. Although there were positive effects with the deposit contract intervention, Study 3 was not designed to evaluate

which behavior mechanism may have been responsible for the behavior change. It is possible that the deposit contract functioned as a discriminative stimulus. For example, the deposit contract might have signaled the availability of positive reinforcement (i.e., earning money) for exercising. The deposit contract may also have functioned as a discriminative stimulus for negative reinforcement in the form of responding (exercise min) to avoid losing their deposit. However, given the package nature of this condition, it is not possible to separate the potential discriminative function of the deposit contract from other contingencies arranged in the current study. For example, participants may have engaged in responding to avoid constructive feedback or to receive positive feedback from the experimenter. More research evaluating specific components independently would allow one to draw conclusions on the mechanisms operating during this condition.

There are many stimuli associated with the competition that may function as discriminative stimuli (e.g., the rings, leaderboard, feedback, etc.). These stimuli associated with the competition may also be conceptualized as gamification. Gamification has been defined as “the use of [video] game design elements in nongame contexts” (Deterding et al., 2011), with common elements including points, leaderboards, achievements, levels, a story or theme, goals, feedback, and competition/ cooperation between players (Hamari et al, 2014). Although the Apple Watch competition does not include an overarching story, it includes many of the other gamification stimuli: goals, feedback, competition or cooperation between players, points, leaderboards, and achievements. These stimuli (e.g., checking the leader board to view the point breakdown, feedback from the Apple Watch, goals of closing the rings, etc.) may function as discriminative stimuli as they may occasion responding previously paired with reinforcement and increase the future probability of those responses (i.e., in this case, physical activity).

Alternatively, these elements, or stimuli, may function as an establishing operation by providing stimuli that visually depict the passage of time since one last exercised. These stimuli may serve to increase motivation to win a competition. More research is needed to determine what makes gamification effective for individuals.

There are several limitations of the current study that deserve comment. First, because participants were required to own an Apple Watch, the generality of the findings to individuals who do not own an Apple Watch remains unclear. To circumvent this limitation, researchers could provide an Apple Watch to participants or evaluate the competition feature on other fitness trackers. Second, the method for pairing participants differed across studies. Second, during Study 1, four (of six total) participants were paired based on points during baseline, rather than exercise min. Because the goal was to increase exercise min, this measure was used in lieu of points for the remaining participants. Third, the current study did not control for adherence to charging and wearing the Apple Watch each day. Overall, most participants charged and wore their watch every day. In Study 2, participant 6 forgot to wear their watch on 5 occasions and forgot to bring their charger while out of town during the second competition, resulting in an early end to that competition and the need to participate in a third competition. During Study 3, participant 2 started a new job that did not allow them to wear a watch. Although this participant indicated that they would like to continue participation in the study, the experimenter needed to send eight reminders to charge and wear their watch. It is possible that these challenges associated with wearing and charging one's watch impacted their results. Future research could consider additional contingencies for charging and wearing the watch when necessary.

Despite these limitations, the present analysis demonstrates the utility of the competition on the Apple Watch in isolation and in combination with additional treatment components. We

recommend that individuals who would like to increase their level of physical activity and own fitness trackers consider using the approaches evaluated in this study. Across the three studies, the competition alone was effective to increase exercise min to performance criteria for six participants ($n = 20$). Competitions are a no-cost social opportunity to engage in more health-related behavior. In Study 2, DRA (\$30 to winner of the competition) was effective to increase exercise min to performance criteria for two participants. DRA could be easily arranged among friends. For example, two people may wager a lunch or dinner between two people as a friendly competition. During Study 2, DOR was added to the competition and effective to increase exercise min to performance criteria for two of the four participants that experienced this condition. DOR could also be easily set up between friends competing, such as sending one another a screen shot the rings or leaderboard across the day. Last, during Study 3, a deposit contract was added to the competition and DOR. This was effective to increase exercise min for all participants that experienced the condition, although only three of five participants met performance criteria. A practical application of a deposit contract may be an arrangement between two friends who are not only competing but would also like to accomplish specific goals (e.g., 3 work out classes a week, go for a 10 min walk, try an at home workout, etc.). To aid in holding one another accountable, they may give one another something to hold onto until they meet their goals (e.g., money or something of personal value). This would be a no cost way to improve health related behavior. Taken together, the current study provides several options for individuals interested in increasing exercise min, or MVPA to lead a healthy lifestyle and avoid preventable disease and health complications.

References

- American Heart Association. (2018, March 31). *Getting Physically Active*.
<https://www.heart.org/en/health-topics/cardiac-rehab/getting-physically-active>
- American Heart Association. (2018, April 18). *American Heart Association Recommendations for Physical Activity in Adults and Kids*. <https://www.heart.org/en/healthy-living/fitness/fitness-basics/aha-recs-for-physical-activity-in-adults>
- Apple. (2019, September 19). *Get the Most Accurate Measurements Using Your Apple Watch*.
<https://support.apple.com/en-us/HT207941>
- Apple. (2020, September 16). *Use the Activity app on Your Apple Watch*.
<https://support.apple.com/en-us/HT204517>
- Apple. (2021, October 4). *Track Your Trends in the Fitness App*. <https://support.apple.com/en-us/HT210343>
- Apple. (2020, December 20). *Monitor Your Heart Rate with the Apple Watch*.
<https://support.apple.com/en-us/HT204666>
- Apple. (2021, December 20). *Share your Activity and compete with friends with your Apple Watch*. <https://support.apple.com/en-us/HT207014>
- Apple. (2022, November 8). *Change Your Goals on Your Apple Watch*.
<https://support.apple.com/en-us/HT212501>
- Centers for Disease Control and Prevention. (2017, February 17). *Glossary*. [NHIS - Adult Physical Activity - Glossary \(cdc.gov\)](https://www.cdc.gov/nchs/nhanes/glossary/physical-activity)
- Centers for Disease Control and Prevention. (2020, November 17). *Social Distancing*.
<https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html>

- Centers for Disease Control and Prevention. (2020, December 6). *Suggestions for Overcoming Physical Activity Barriers*. <https://www.cdc.gov/physicalactivity/basics/adding-pa/barriers.html>
- Centers for Disease Control and Prevention. (2020, May 14). *Physical Activity Prevents Disease*. <https://www.cdc.gov/chronicdisease/resources/infographic/physicalactivity.htm#:~:text=Regular%20physical%20activity%20helps%20improve,depression%20and%20anxiety%2C%20and%20dementia.>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis*. Columbus, Ohio, Merrill Pub. Co.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining “gamification.” *In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011* (pp. 9–15). <https://doi.org/10.1145/2181037.2181040>
- Donlin Washington, W., McMullen, D., & Devoto, A. (2016). A matched deposit contract intervention to increase physical activity in underactive and sedentary adults. *Translational issues in Psychological Science, 2*(2), 101-114. <https://doi.org/10.1037/tps0000069>
- Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills: recommendations for instructors. *Behavior analysis in practice, 6*(1), 56–75. <https://doi.org/10.1007/BF03391791>
- Jarvis, B. P., & Dallery, J. (2017). Internet-based self-tailored deposit contracts to promote smoking reduction and abstinence. *Journal of applied behavior analysis, 50*(2), 189–205. <https://doi.org/10.1002/jaba.377>

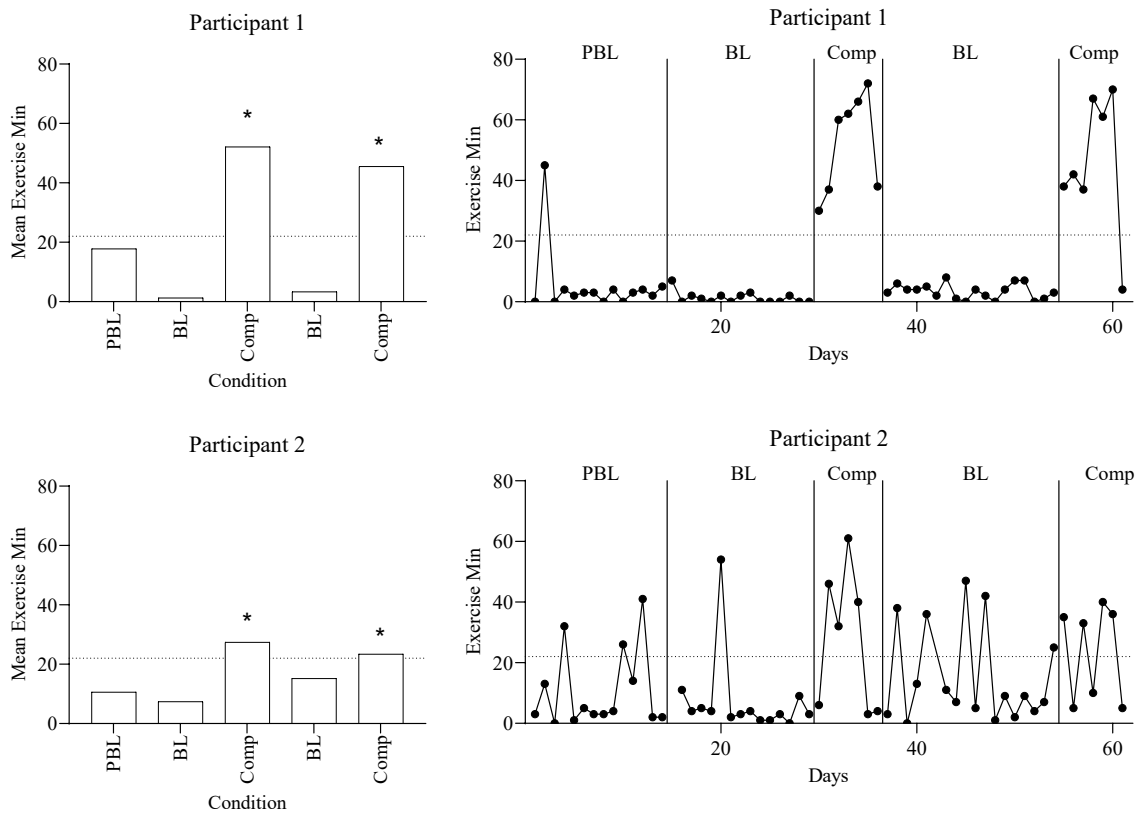
- Krentz, H., Miltenberger, R., & Valbuena, D. (2016). Using token reinforcement to increase walking for adults with intellectual disabilities. *Journal of Applied Behavior Analysis*, 49(4), 745-750. <https://doi.org/10.1002/jaba.326>
- Lillie, M. A., & Tiger, J. H. (2019). Acquisition and generative responding following print-to-braille construction response training with sighted learners. *Journal of Applied Behavior Analysis*, 52(1), 286-298. <https://doi.org/10.1002/jaba.516>
- Normand, M. P., & Burji, C. (2020). Using the Step it UP! game to increase physical activity during physical-education classes. *Journal of applied behavior analysis*, 53(2), 1071-1079. <https://doi.org/10.1002/jaba.624>
- Stedman-Falls, L. M., & Dallery, J. (2020). Technology-based versus in-person deposit contract treatments for promoting physical activity. *Journal of applied behavior analysis*, 53(4), 1904–1921. <https://doi.org/10.1002/jaba.776>
- Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the Physical Activity Readiness Questionnaire (PAR-Q). *Canadian Journal of Sport Sciences*, 17(4), 338–345.
- Valbuena, D., Miltenberger, R., & Solley, E. (2015). Evaluating an Internet-based program and a behavioral coach for increasing physical activity. *Behavior Analysis: Research and Practice*, 15(2), 122-138. <http://dx.doi.org/10.1037/bar0000013>
- Washington, W. D., Banna, K. M., & Gibson, A. L. (2014). Preliminary efficacy of prize-based contingency management to increase activity levels in healthy adults. *Journal of applied behavior analysis*, 47(2), 231–245. <https://doi.org/10.1002/jaba.119>
- World Health Organization. November 16, 2020. *Physical Activity*. <https://www.who.int/news-room/fact-sheets/detail/physical->

[activity#:~:text=Globally%2C%2028%25%20of%20adults%20aged,intensity%20physical%20activity%20per%20week](#)

\

Figure 1.

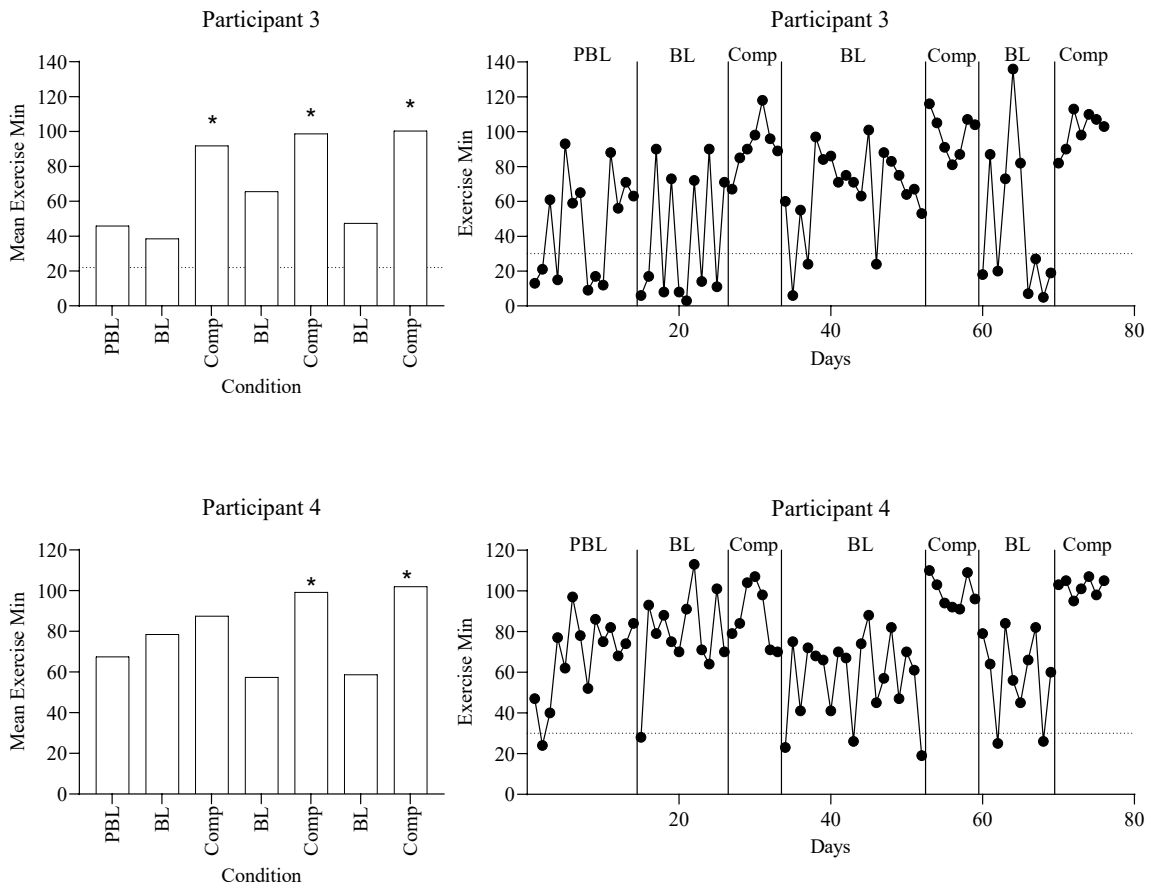
Treatment Analysis Results for Participants 1 and 2 (Study 1)



Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), and the competition (Comp).

Figure 2.

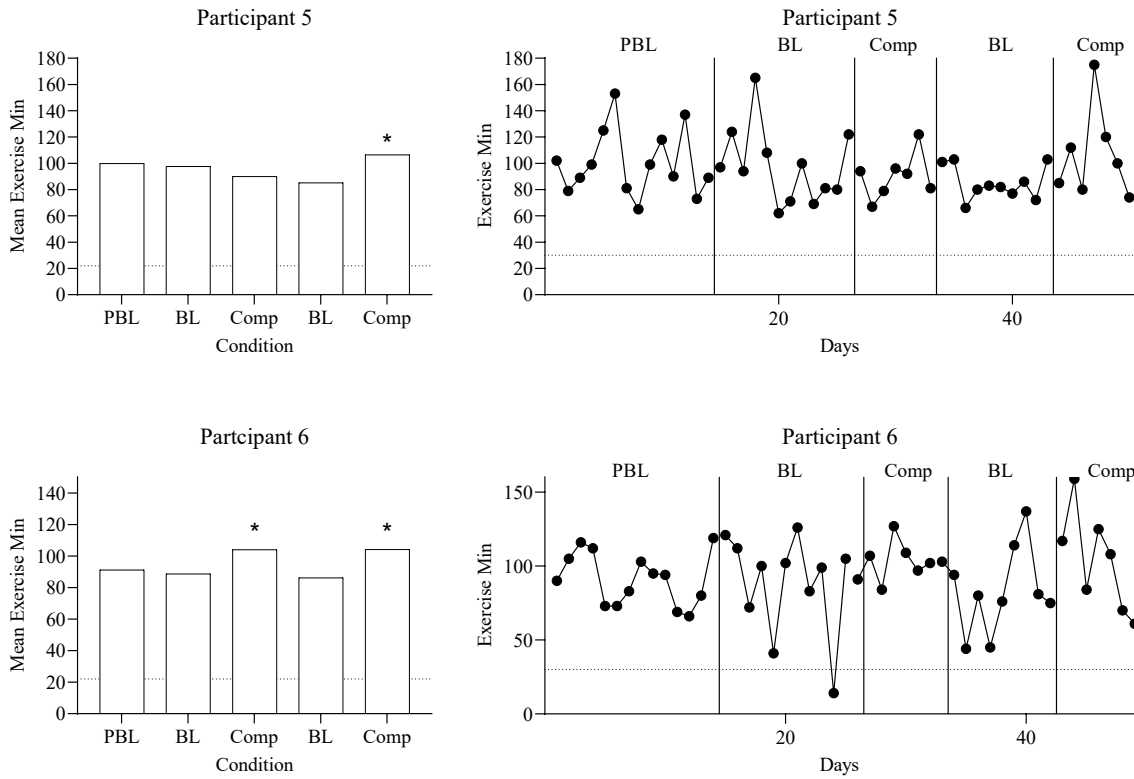
Treatment Analysis Results for Participants 3 and 4 (Study 1)



Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), and the competition (Comp).

Figure 3.

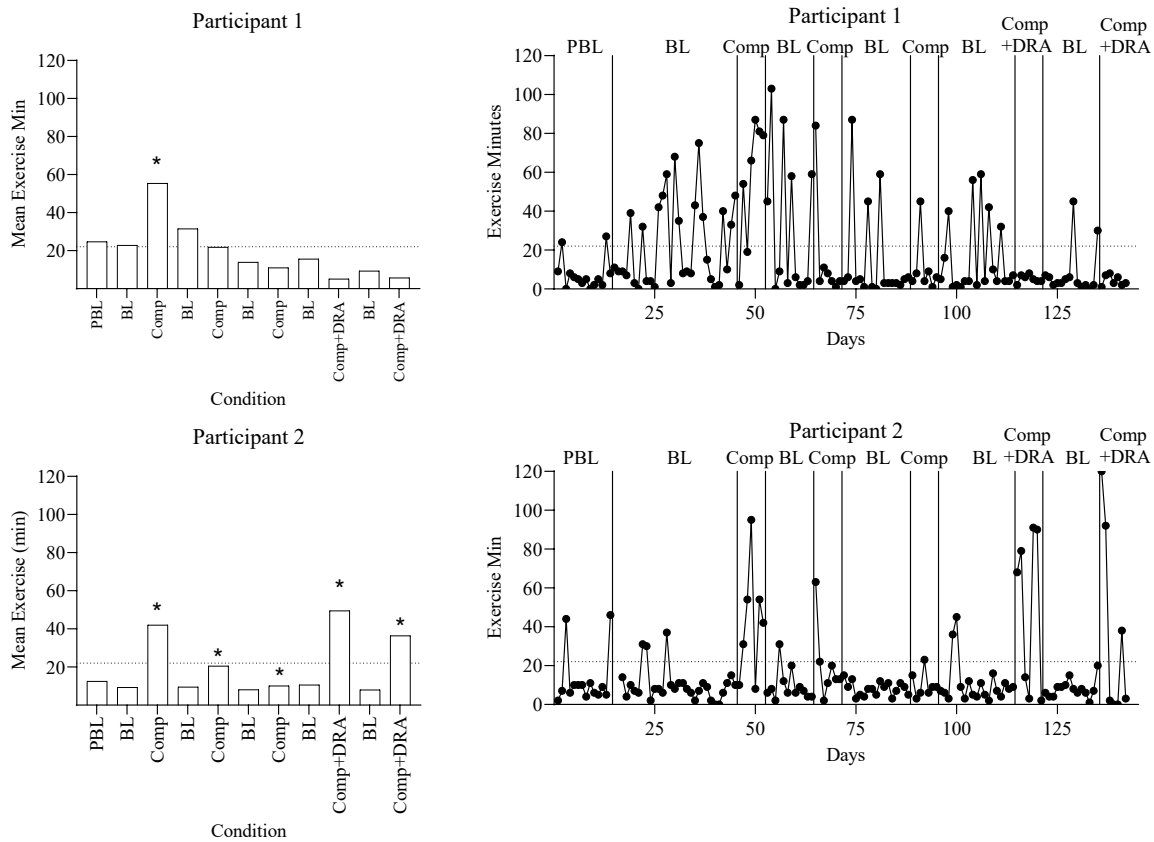
Treatment Analysis Results for Participants 5 and 6 (Study 1)



Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), and the competition (Comp).

Figure 4.

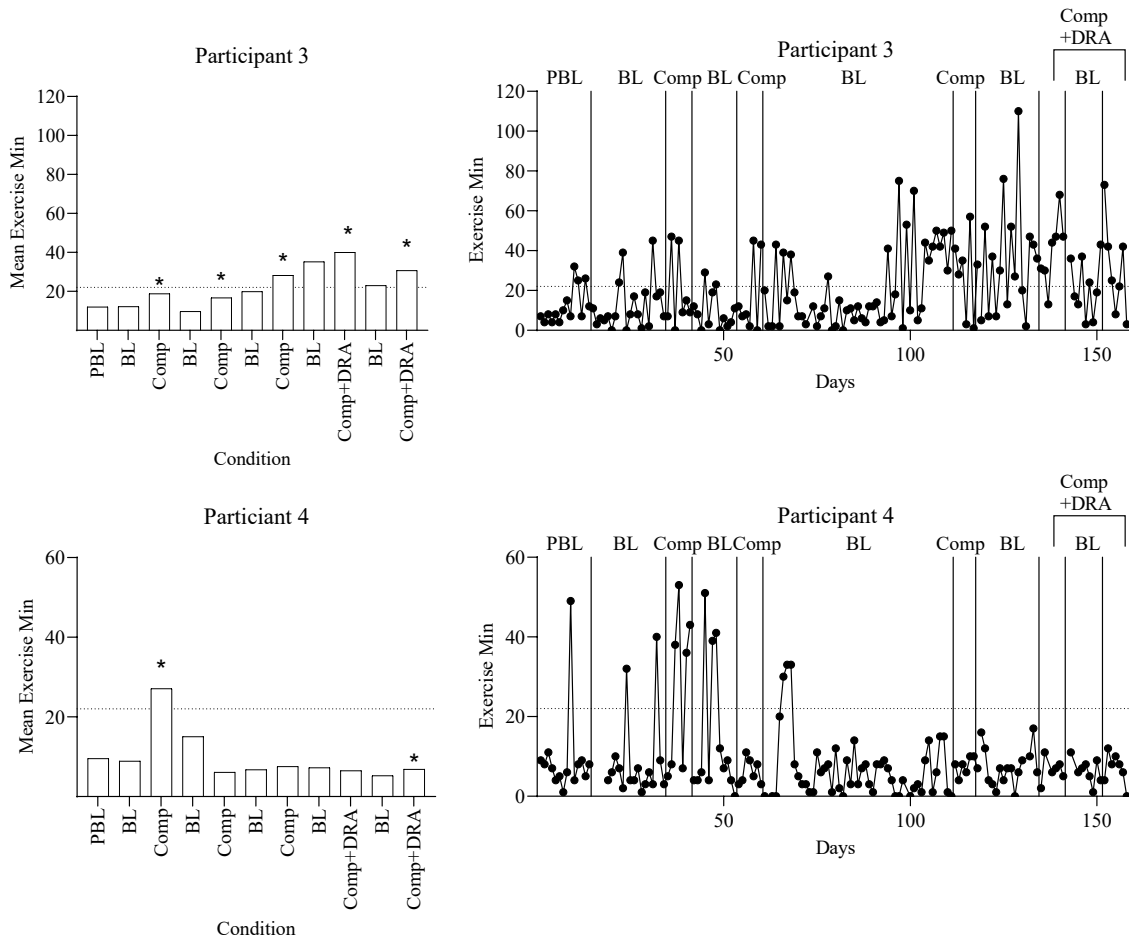
Treatment Analysis Results for Participants 1 and 2 (Study 2)



Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), the competition (Comp), and competition plus DRA (Comp + DRA).

Figure 5.

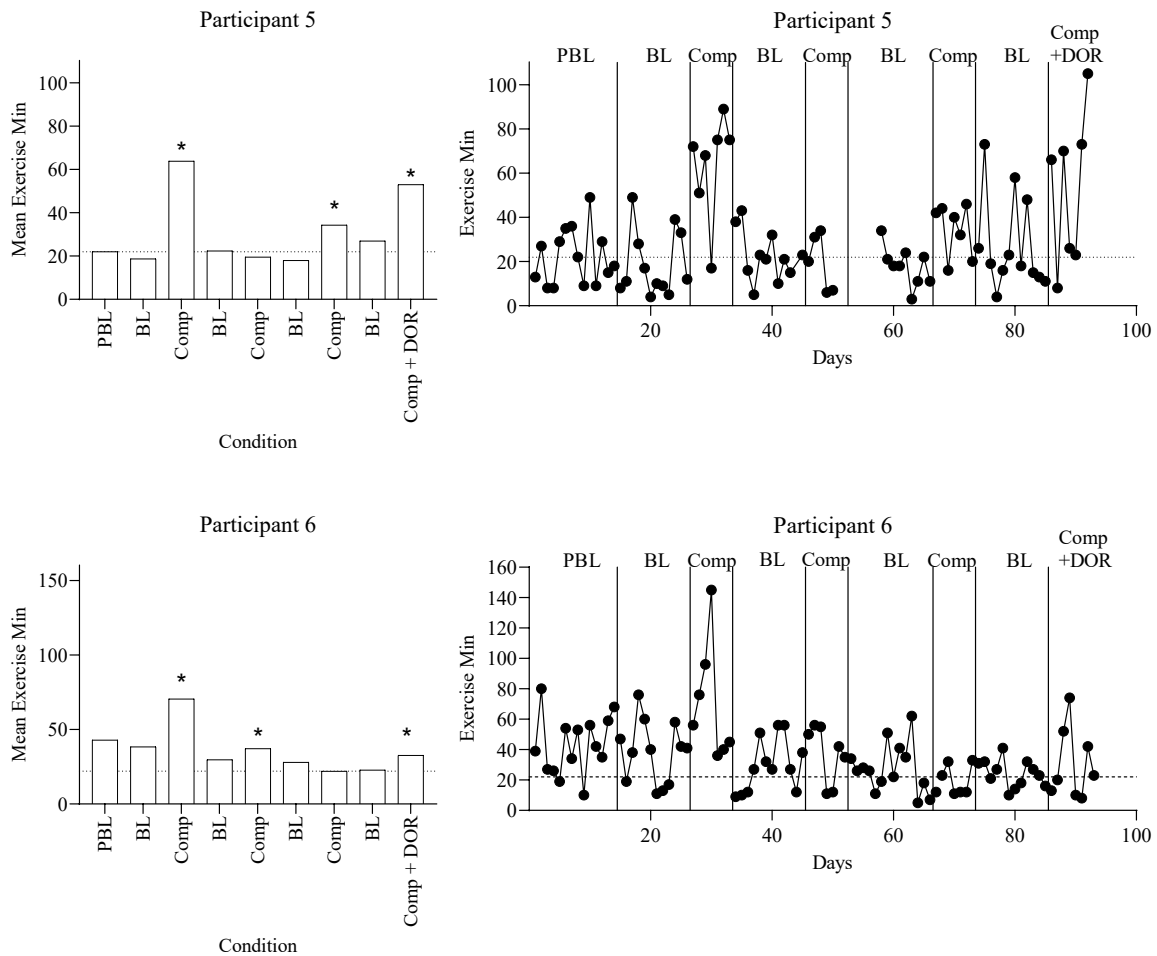
Treatment Analysis Results for Participants 3 and 4 (Study 2)



Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), the competition (Comp), and competition plus DRA (Comp + DRA).

Figure 6.

Treatment Analysis Results for Participants 5 and 6 (Study 2)

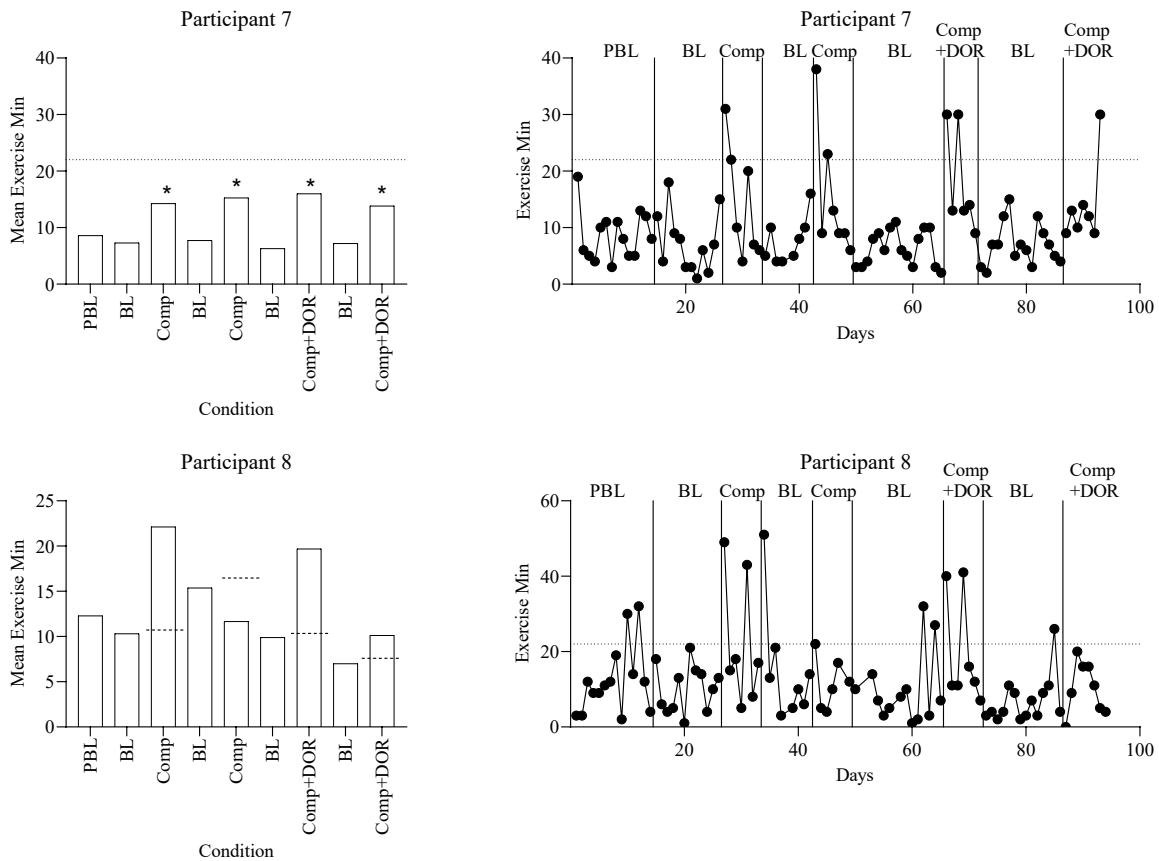


Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day.

The conditions evaluated included pre-baseline (PBL), baseline (BL), the competition (Comp), and competition plus DOR (Comp + DOR).

Figure 7.

Treatment Analysis Results for Participants 7 and 8 (Study 2)

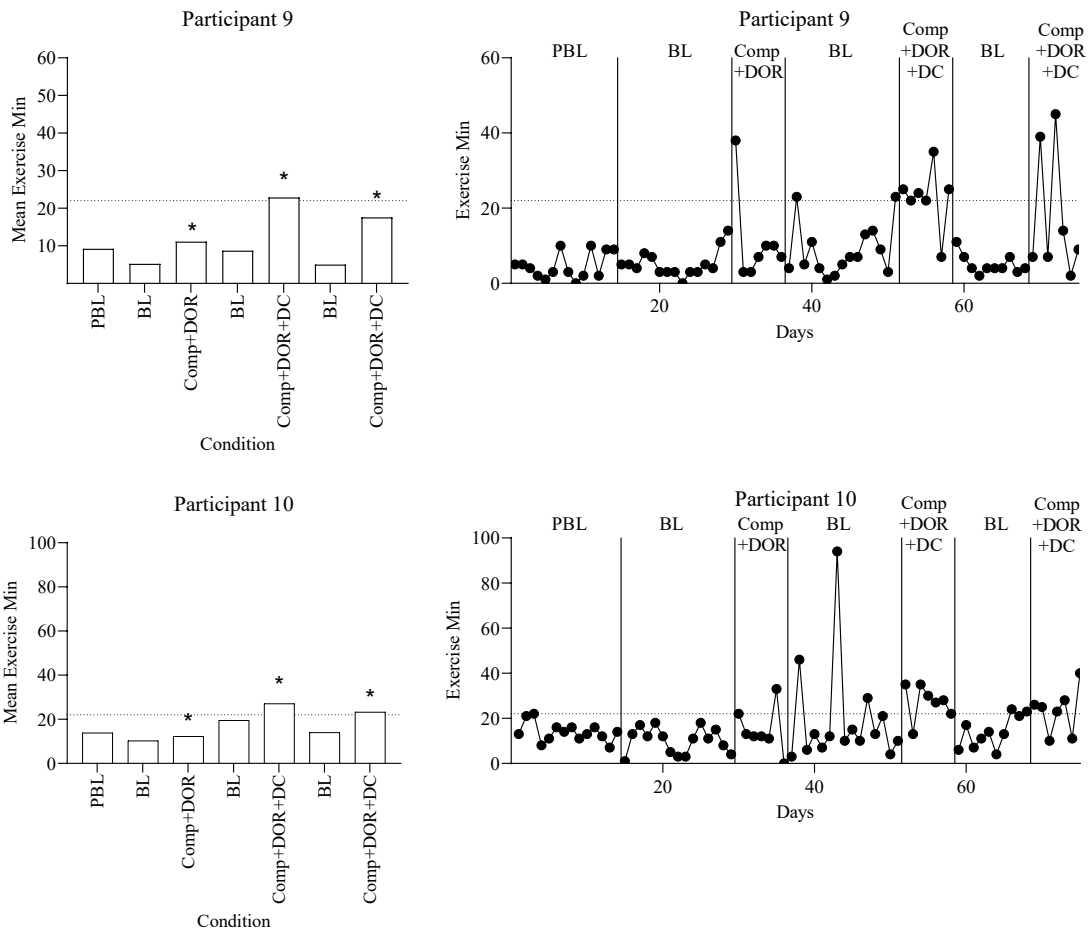


Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day.

The conditions evaluated included pre-baseline (PBL), baseline (BL), the competition (Comp), and competition plus DOR (Comp + DOR).

Figure 8.

Treatment Analysis Results for Participants 9 and 10 (Study 2)

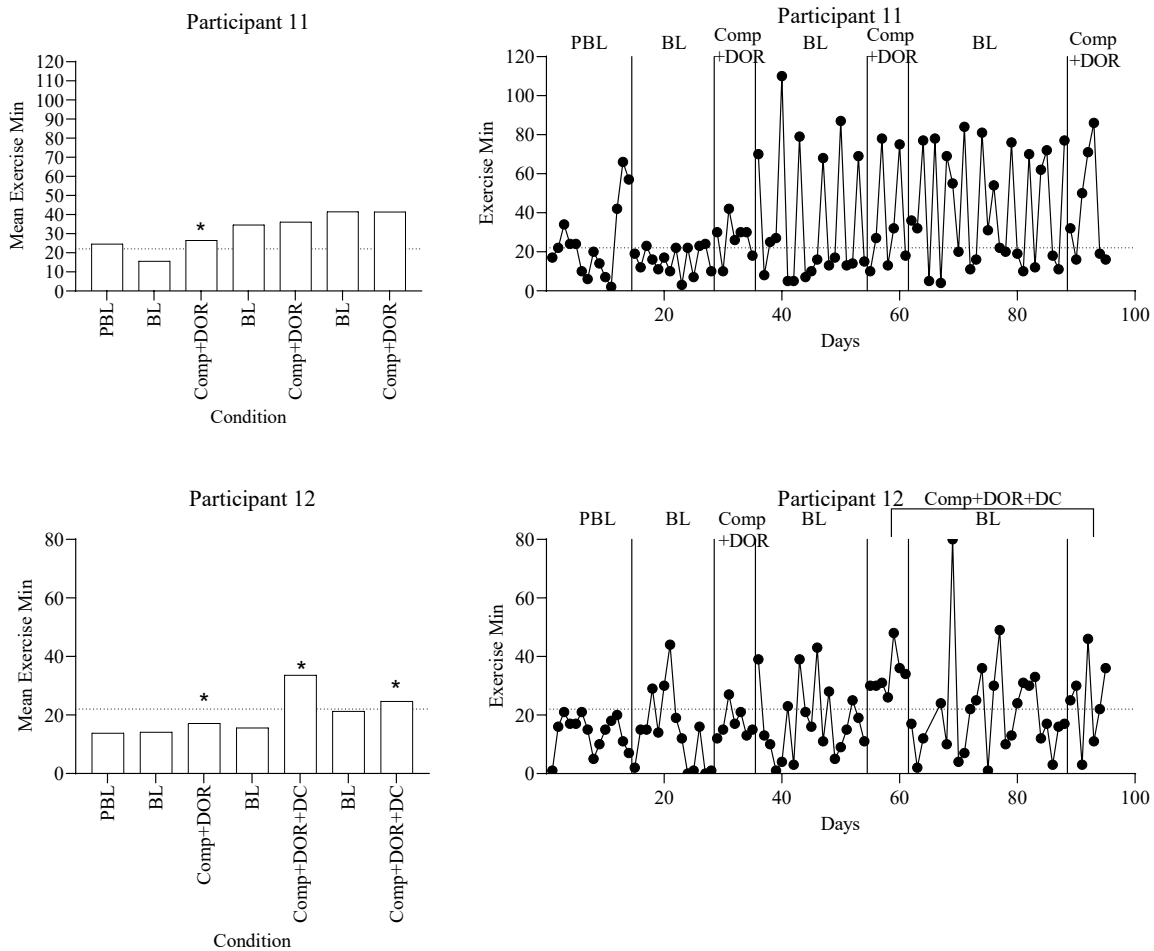


Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for

each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), competition plus DOR (Comp + DOR), and competition plus DOR plus deposit contract (Comp + DOR + DC).

Figure 9.

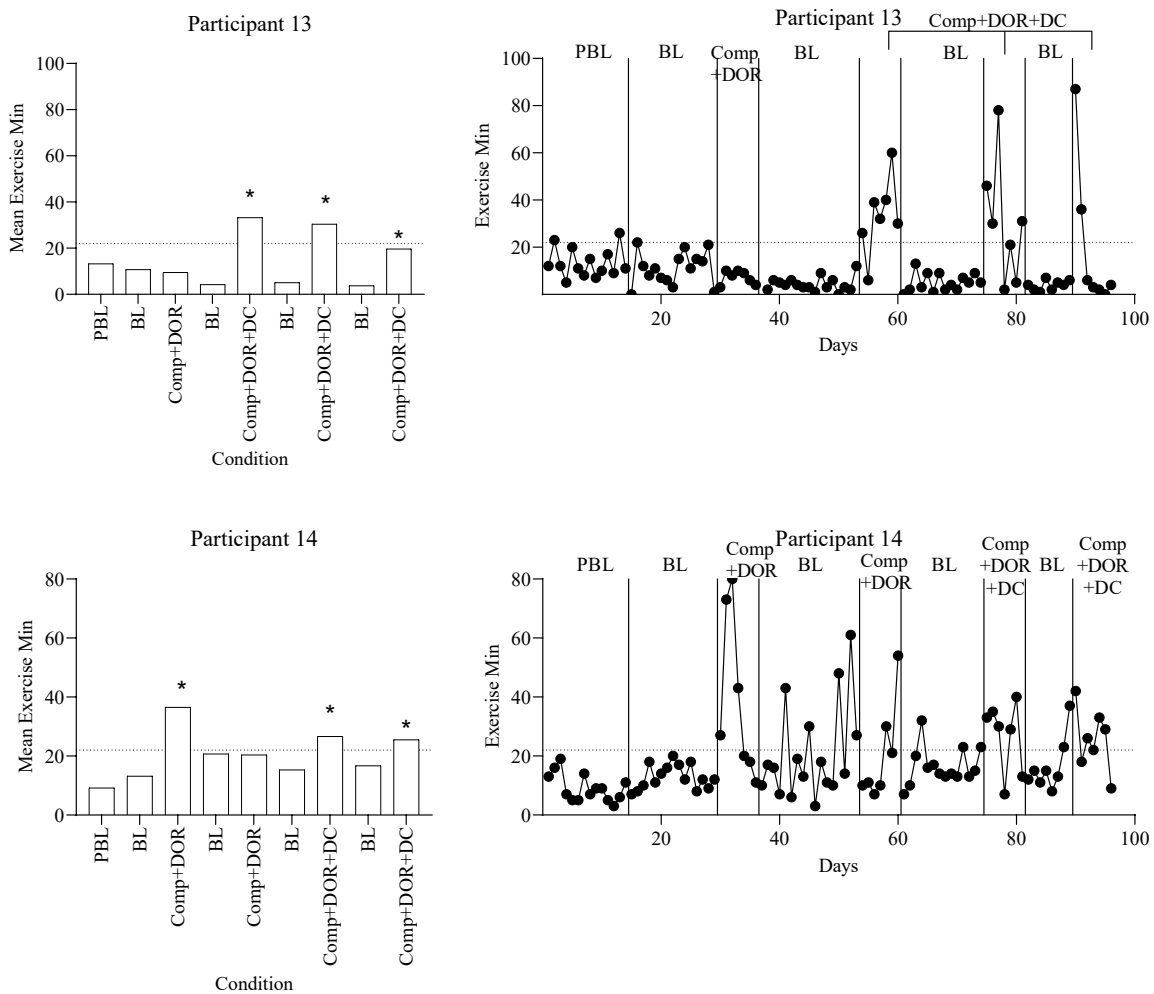
Treatment Analysis Results for Participants 11 and 12 (Study 2)



Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), competition plus DOR (Comp + DOR), and competition plus DOR plus deposit contract (Comp + DOR + DC).

Figure 10.

Treatment Analysis Results for Participants 13 and 14 (Study 2)



Note. The graphs on the left panel depict mean exercise min per condition, the dashed line indicates the 22-exercise min a day, and the asterisks indicate when participants met the percentage change performance criteria. The graphs on the right panel depict exercise min for each day within each condition and a dashed line to indicate the target 22-exercise min a day. The conditions evaluated included pre-baseline (PBL), baseline (BL), competition plus DOR (Comp + DOR), and competition plus DOR plus deposit contract (Comp + DOR + DC).