

Chapman University

Chapman University Digital Commons

Psychology Faculty Articles and Research

Psychology

3-5-2020

Variable Magnitude and Frequency Financial Reinforcement is Effective at Increasing Adults' Free-Living Physical Activity

Vincent Berardi

Chapman University, berardi@chapman.edu

Melbourne Hovell

San Diego State University

Jane C. Hurley

Arizona State University

Christine B. Phillips

Arizona State University

John Belletierre

University of California, San Diego

See next page for additional authors

Follow this and additional works at: https://digitalcommons.chapman.edu/psychology_articles



Part of the [Behavioral Disciplines and Activities Commons](#), [Behavior and Behavior Mechanisms Commons](#), [Health Psychology Commons](#), and the [Other Psychiatry and Psychology Commons](#)

Recommended Citation

Berardi, V., Hovell, M., Hurley, J.C., Phillips, C. B., Belletierre, J., Todd, M., & Adams, M. A. (2020). Variable magnitude and frequency financial reinforcement is effective at increasing adults' free-living physical activity. *Perspectives on Behavior Science*, 43, 515-538. <https://doi.org/10.1007/s40614-019-00241-y>

This Article is brought to you for free and open access by the Psychology at Chapman University Digital Commons. It has been accepted for inclusion in Psychology Faculty Articles and Research by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Variable Magnitude and Frequency Financial Reinforcement is Effective at Increasing Adults' Free-Living Physical Activity

Comments

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in *Perspectives on Behavior Science*, volume 43, in 2020 following peer review. The final publication may differ and is available at Springer via <https://doi.org/10.1007/s40614-019-00241-y>

[Please click here for a read-only version of the final, published version.](#)

Copyright

Association for Behavior Analysis International

Authors

Vincent Berardi, Melbourne Hovell, Jane C. Hurley, Christine B. Phillips, John Belletierre, Michael Todd, and Marc A. Adams

Variable Magnitude and Frequency Financial Reinforcement is Effective at Increasing Adults' Free-Living Physical Activity

Keywords: reinforcement schedules; physical activity promotion; adaptive goals; accelerometer; health technology

Abstract

Financial rewards can increase health behaviors, but little research has quantified the effects of different reinforcement schedules on this process. This analysis compares the average moderate-to-vigorous physical activity (MVPA) associated with six distinct positive reinforcement schedules implemented within a physical activity promotion clinical trial. In this trial, participants ($N=512$) wore an accelerometer for one year and were prescribed one of two types of MVPA goals: a static 30-minute goal or an adaptive goal based on the MVPA produced over the previous nine days. As participants met goals, they transitioned through a sequence of reinforcement stages, beginning with a continuous-fixed magnitude (CRF-FM), then CRF-variable magnitude (CRF-VM), followed by a series of variable ratio-VM (VR-VM) schedules. The average accumulation of MVPA bouts over the last 24 days of each stage was compared to each other. Average MVPA during stage transitions were also examined. The results indicated that immediate reinforcement resulted in more MVPA relative to a comparison group and that the relative effectiveness of adaptive versus static goals was dependent on the magnitude of daily MVPA goals. Schedule effects were qualitatively different for individuals who frequently met their daily goals (Large Intervention Effect subgroup) versus those who did not (Small Intervention Effect subgroup). For the Large Intervention Effect group, the CRF-VM schedule produced the most MVPA, particularly within the adaptive goal condition, with increases observed immediately upon encountering this schedule. In contrast, the CRF-FM schedule produced small amounts of MVPA. This pattern was reversed for the Small Intervention Effect subgroup, where the most MVPA was associated with the CRF-FM stage. Future interventions should focus on discriminating small versus large intervention effects as quickly as possible so that the optimal reinforcement schedule can be used.

Introduction

Well-documented benefits of regular moderate-to-vigorous physical activity (MVPA) include improvements to cardiorespiratory fitness, body composition, bone health, mental health, and reduced risks for cardiovascular disease and several cancers (U.S. Department of Health and Human Services, 2008, 2018). Nevertheless, only a small percentage (<8%) of US adults meet federally-recommended MVPA guidelines (Troiano et al., 2008), and there has been insufficient change in population levels of activity over the past two decades, even in the context of significant investments in public health (Brownson, Boehmer, & Luke, 2005; Kruger, Ham, & Kohl III, 2005; Ward, Clarke, Nugent, & Schiller, 2015). To address this public health problem, a large number of studies have tested interventions to increase PA using a wide range of behavior change strategies and theoretical models. A meta-analysis of individual-level interventions designed to increase PA among healthy adults (N=99,001, 358 papers) found an overall mean difference of only 2.1 min/day (95% CI -1.57 to 5.76) favoring PA interventions over comparison groups (Conn, Hafdahl, & Mehr, 2011). This suggests a need for more potent interventions.

Multiple behavior-change theories recognize positive reinforcement as a critical component when encouraging healthy habits (Bandura, 1997; Michie et al., 2013; Prochaska, Redding, & Evers. KE, 2015; Skinner, 1953). As result, interventions often make use of contingency management (CM) procedures that monitor a behavior and provide a reinforcer when some criterion is met (Donaldson & Normand, 2009; McPherson et al., 2018; Washington, Banna, & Gibson, 2014; J. Weinstock, Barry, & Petry, 2008). Similarly, many PA interventions have incorporated financial incentives into their design (Barte & Wendel-Vos, 2017; Irons, Pope, Pierce, Van Patten, & Jarvis, 2013; Mitchell et al., 2013; Patel, Normand, & Kohn, 2019), with wide variation in effectiveness (Giles, Robalino, McColl, Sniehotta, & Adams, 2014) potentially due, in part, to deficiencies in the models of behavior change that have guided these studies (J. Adams, Giles, McColl, & Sniehotta, 2014; Giles et al., 2016).

Behavior continuously varies and sources of variation can be internal or external (Hovell, Wahlgren, Adams, & Wahlgren, 2009; Rachlin, 1994, 2000; Riley et al., 2011). Yet most studies measure MVPA infrequently by self-report, which has been shown to be inconsistent with objective measures (Skender et al., 2016), and/or used cognitive models that do not fully account for or harness sources that explain the dynamic nature of day-to-day MVPA (M. A. Adams et al., 2013; Rowlands et al., 2015). Riley et al. (2011) have argued that the next generation of health behavior interventions using near real-time measures of free-living behaviors will reveal sources of variation and must be guided by dynamic, regulatory, and adaptive models of behavior change that consider an *individual's* current and past behavior and context. Operant theory meets these criteria and, as a result, has been proposed as a unifying approach to health-promotion science (Biglan, 2003).

Many health behavior studies of free-living individuals that use reinforcement fail to appreciate the operant foundations of this approach. Even when operant characteristics are recognized, dimensions of reinforcement that have been established within basic research are often neglected, which may reduce intervention effectiveness. This can lead interventions to have to suboptimal features such as a long delay between response and consequence, which is known to weaken reinforcement effects (Chung & Herrnstein, 1967); the use of limited, familiar reinforcers, which are less effective than novel stimuli (Berlyne, Koenig, & Hirota, 1966); a focus on loss aversion from behavioral-economic cognitive models, despite evidence that positive reinforcement is preferred and more acceptable to participants (Halpern et al., 2015); a failure to explicitly design for resistance to extinction (S. Weinstock, 1954); or some combination of the above.

One feature of contingent reinforcement that is often overlooked in the physical activity literature is the reinforcement schedule, which can be characterized according to steady-state patterns of behavioral response and differential rates of extinction (Ferster & Skinner, 1957). For example, variable ratio schedules, which provide reinforcement after a known, but unpredictable average number of

responses, produce high frequency of responding relative to continuous reinforcement (Ferster & Skinner, 1957). Within human health interventions, reinforcement schedules have been identified as factors likely affecting the adoptability and sustainability of behavior adoption (Johnston, 2016), yet we are aware of only a handful of studies that have systematically examined schedules for PA. De Luca and Hoborn (1990, 1992) conducted a series of studies exploring the effects of interval and variable ratio schedules on adolescent boys' PA. Other studies investigated the effects of variable ratio schedules on obese children's selection of sedentary versus vigorous activity (Epstein, Smith, Vara, & Rodefer, 1991; Epstein et al., 1995). While these studies demonstrated the utility of variable ratio schedules in generating healthy behavior, they were performed in clinical settings with children and the extent to which results generalize to other populations and function in real-world settings where they must compete with existing, and often counterproductive, contingencies remains unknown.

The Walking Intervention Through Texting (WalkIT) Arizona study provided a PA intervention designed to address several of the deficiencies of typical studies that implement PA goals and financial incentives. This year-long, free-living intervention used near real-time assessments of MVPA and, in a process consistent with operant theory, implemented varying financial reinforcement schedules. The study aimed to demonstrate increases in MVPA by contrasting a group that received immediate financial reinforcement for meeting PA goals to a group that received financial rewards on a 60-day interval schedule that was not contingent on PA (henceforth, called the *comparison* group). Participants in the immediate reinforcement group progressed through one to six different "reinforcement stages," with stages defined by different schedules of reinforcement. From a theoretical standpoint, these schedules were conceptualized to initially shape MVPA adoption and then to develop resistance to extinction via intermittent reinforcement, but their functionality over a long-term intervention for free-living behavior is unknown. Therefore, the purpose of the analysis that follows is to explore the differential effects of these reinforcement schedules on participants' MVPA throughout their enrollment in the study. More

specifically, we explored: 1) whether the initial presentation of immediate reinforcement vs. comparison increased MVPA relative to baseline, 2) whether certain reinforcement stages resulted in more MVPA than others, 3) how quickly the effects of reinforcement stages were observed, and 4) whether temporal effects throughout exposure to a single reinforcement stage were present.

Methods

WalkIT Arizona Design

The WalkIT Arizona study aimed to increase MVPA among 512 insufficiently active men and women from Maricopa County, Arizona. Enrolled participants were asked to wear an ActiGraph GT9X wrist-worn accelerometer daily during awake hours for one year and to sync (i.e., upload) data from their accelerometer each day to WalkIT Arizona project servers. A MVPA goal was provided to participants each day and, after syncing, participants were informed of whether this goal was met, with select participants receiving financial reinforcement for achieving their goal. An automated, cloud-based system was designed for this study and online 24 hours/day, 365 days per year to receive and process accelerometer data, determine whether goals were met, calculate financial rewards, transmit text message feedback to participants, and send e-gift cards, when appropriate. WalkIT Arizona was pre-registered at *clinicaltrials.gov* (NCT02717663) and a full description of its methodology and sample characteristics is available in Adams et al (2019). A brief description is provided below.

The study aimed to recruit an equal number of men and women aged 18 to 60 years from four neighborhood quadrants characterized by high versus low household income and high versus low walkability. Recruitment primarily occurred through Facebook, because it allowed for demographic and geographic targeting of prospective participants on an ongoing basis. Participants were initially enrolled in a 10-day baseline phase as WalkIT Arizona's inclusion criteria was assessed as MVPA was passively recorded using a blinded accelerometer. Some participants had baseline phases longer than 10 days due to issues such as accelerometer or mobile app problems, non-adherence to accelerometer wear protocol,

or illness. After the baseline, each participant meeting the inclusion criteria was randomized into one of four groups, stratified as a 2 x 2 factorial design that tested two PA goal conditions (Static versus Adaptive) by two financial reinforcement conditions (Immediate versus Comparison). These conditions were operationalized as follows.

- *Static PA Goals:* Set to 30 minutes of MVPA per day and did not change over the study.
- *Adaptive PA Goals:* Set to equal the 60th rank-order percentile based on moving window of MVPA minutes accrued over the previous nine days and tested in prior research for PA (M. A. Adams, 2009; M. A. Adams et al., 2017, 2013). More specifically, the daily MVPA for a single participant was recorded over the previous nine days and rank-ordered from lowest to highest. The MVPA value associated with the 60th percentile was calculated via Galbicka's (1994) equation $k = (m + 1)(1 - w)$; where k is the rank order, m is the number of prior observations in the window, and w is the desired percentile. The MVPA value with rank order k was set as the next day's goal, which the participant had to meet or exceed to earn financial reinforcement. Each new day's MVPA observation replaced the oldest observation in the moving window, meaning that adaptive goals could change daily.
- *Immediate Reinforcement:* Earned points when accumulated daily MVPA bout-minutes met or exceeded their daily goal and they synced their device by noon on the following day. Upon meeting these two criteria, participants were immediately provided with points that could be exchanged for cash-equivalent gift cards once 500 points (i.e., \$5.00) had been accumulated. The noon next-day deadline for syncing encouraged participants to sync frequently and ensured that an undue amount of time did not elapse between the engagement in physical activity and the presentation of a reinforcer. As described below, reinforcement magnitude and frequency

varied as participants progressed through a series of reinforcement schedules labeled “reinforcement stages.”

- *Comparison Treatment:* These individuals received financial reinforcement on a 60-day interval schedule, after successfully wearing their accelerometer for at least 10 hours on 4 out of the last 7 days and syncing the device. If a participant did not meet these criteria once the 60 days had elapsed, a reminder was sent via text message. This group did not have a contingency for meeting MVPA goals and received up to \$265 in cash-equivalent gift cards on an escalating magnitude basis: \$15 in month 2 (M2), \$30 in M4, \$50 in M6, \$75 in M8, and \$95 in M10.

Regardless of reinforcement group or goal type, after successfully meeting a daily goal and syncing their accelerometer, a feedback text message was provided with praise and the next goal (e.g., “Cheers, James! Goal met! 63 min yesterday. Goal for 7/1 is 35 min). For immediate reinforcement participants, the text also included the number of points earned and current point balance (e.g. “Reward points = 100! Balance is 400 points.”). Participants were able to choose among several cash-equivalent gift card options (Amazon, Target, Walmart, Starbucks, etc.) for reinforcement and could switch their selection at any time by texting or emailing WalkIT Arizona staff with the desired change. We sent reminders about this option periodically.

Each participant in the immediate reinforcement group was subject to up to six different reinforcement schedules, which were presented in a series of stages (Table 1). The speed at which a participant progressed through this series and the number of stages encountered was dictated by the frequency with which daily MVPA goals were met. Participants progressed out of a given stage by meeting 24 daily MVPA goals, which corresponded to the receipt of \$24 in monetary compensation. Upon reaching a new stage, the 24-goal requirement was restarted and another \$24 in reinforcement could be earned. Participants were blind to the specifics of the reinforcement schedules/stages and were only told that they could earn points each day they met their physical activity goal at 1 point = 1 cent, with the amount

potentially varying from 0 to 500 daily points. Prior to meeting their goal, participants were unaware of the amount that could be earned for the day.

Individuals began the study with a continuous reinforcement (CRF) schedule where every daily goal met earned a fixed magnitude (FM) \$1.00 reward (Stage 1). Stage 2 was characterized by CRF with variable reinforcement magnitudes (VM), ranging from \$0.25 to \$2.50. After Stage 2, participants progressed through several variable ratio (VR), VM schedules. Through Stage 6, these schedules were unique, but in Stages 7 through 9, the schedules regressed backwards to those used in Stages 5 through 3, respectively. In Stage 10 and beyond, the schedule from Stage 3 was repeated until the end of the intervention, which ensured that a sufficiently rich schedule of intermittent reinforcement was encountered during the year-long study. The determination of which successful responses would be reinforced was selected at random for each participant prior to their enrollment.¹ This process allowed the reinforcement rate in each stage to be precisely controlled. For instance, in Stage 3, a VR 1.09 schedule was used, which required 22 of the 24 days with a daily goal met to be reinforced. Participant 1 might be reinforced for all but the 3rd and the 19th goal met while Participant 2 could be reinforced for all but the 8th and the 12th goal met, and so forth. This process produced the unpredictability in reinforcement presentation that is characteristic of VR schedules.

	<u>Stage</u>	<u>Description</u>	<u>Reward Ratio</u>	<u>Reward Range per Met Goal</u>	<u>Max \$ Earned in Stage</u>
	Baseline	Prior to Randomization	--	--	--
24 Goals Met	Stage 1	Continuous, Fixed Magnitude	CRF-FM	\$1.00	\$24
	Stage 2	Continuous, Variable Magnitude	CRF-VM	\$0.25-\$2.50	\$24

¹ This protocol does not mirror the traditional deployment of a VRx schedule, where each target response has an independent $1/x$ probability of being reinforced. The randomness inherent in this probabilistic approach can lead to short-term deviations between x and the ratio of reinforced target responses, especially when considering only a small number of responses (24) as was the case for this study. Other studies have used similar *a priori* approaches to construct VR schedules (e.g. Lerman, Iwata, Shore, & Kahng (1996) and Ferster & Skinner (1957, Figures 471-473).

<i>Required to Advance Stage</i>	Stage 3	Variable Ratio, Variable Magnitude	VR 1.09-VM	\$0.50-\$2.50	\$24
	Stage 4	Variable Ratio, Variable Magnitude	VR 1.20-VM	\$0.75-\$2.50	\$24
	Stage 5	Variable Ratio, Variable Magnitude	VR 1.33-VM	\$0.75-\$3.50	\$24
	Stage 6	Variable Ratio, Variable Magnitude	VR 1.50-VM	\$0.75-\$5.00	\$24
	Stage 7	Variable Ratio, Variable Magnitude	VR 1.33-VM	\$0.75-\$3.50	\$24
	Stage 8	Variable Ratio, Variable Magnitude	VR 1.20-VM	\$0.75-\$2.50	\$24
	Stage 9 +	Variable Ratio, Variable Magnitude	VR 1.09-VM	\$0.50-\$2.50	\$24

Table 1: Description of reinforcement schedules stages implemented within the immediate reinforcement group in WalkIT.

Accelerometer Outcome Measure and MVPA Processing

Free-living MVPA was measured by wrist-worn accelerometry (GT9X Link, ActiGraph, LCC, Pensacola, FL, USA) from which MVPA bout-minutes were scored by the automated system mentioned above. An MVPA bout-minute designation was determined for each minute of wear by a binary classification using the following two criteria: 1) whether vector-magnitude per minute (i.e., square root of the sum of squared counts from the accelerometer’s 3 axes) exceeded an individualized “cut point” threshold, and 2) whether the individual’s native ambulatory stepping behavior (i.e., steps per minute) exceeded a 30-step threshold. As fully described in Adams et al. (2019), the vector magnitude cut points were individually determined for each participant via a laboratory-based walking protocol at baseline that approximated well-established procedures (Freedson, Melanson, & Sirard, 1998), but was adapted to determine a personalized cut point threshold for moderate-to-vigorous intensity walking unique to each participant. The onset of an MVPA bout occurred once both the VM and step criteria occurred for a minimum of three non-contiguous minutes out of a moving 5-minute window, and the offset of a bout occurred once these MVPA criteria were not met for 3 contiguous minutes of a moving 5-minute window. The minimum detectable MVPA bout by this approach was 3 minutes (see Adams, et al. (2019) for full details), which is similar to commercial wearable wrist-worn activity monitors. Participants were informed of this bout criteria. The primary outcome measure used for in following analyses was the cumulative number of MVPA bout-minutes across the baseline and various reinforcement stages.

Data Preparation and Sample Description

Valid wear days were considered those on which participants recorded at least ten hours of accelerometer wear time or met their physical activity goal for the day. Participants who did not record at least 48 valid days did not have an opportunity to progress into a second stage; therefore, they were eliminated from all analyses. For a given participant, data from any reinforcement stage in which they did not have a minimum of 24 observations were eliminated since there was not an opportunity to complete and advance out of this stage. This ancillary sample consisted of 501 participants with a total of 167,443 daily observations, with an average of 189.7 days of observation (including baseline) per participant. In total, 127 participants were in the adaptive goal/immediate reinforcement group, 124 participants were in the static goal/immediate reinforcement group, 126 participants were in the adaptive goal/comparison group, and 124 participants were in the static goal/comparison group. The demographics of the analytic sample are provided in Table 2.

	Overall Sample (n=501)	Immediate Reinforcement (n=251)	Comparison (n=250)
Age, mean (SD)	45.8 (8.9)	46.7 (8.5)	44.9 (9.1)
Gender, n (%)			
Female	191 (60.3)	83 (56.1)	108 (63.9)
Male	126 (39.7)	65 (43.9)	61 (36.1)
Ethnicity, n (%)			
White, non-Hispanic	273 (86.1)	129 (87.2)	144 (85.2)
Non-white	27 (8.5)	12 (8.1)	15 (8.9)
Prefer not to answer	17 (5.4)	7 (4.7)	10 (5.9)
BMI, mean (SD)	32.8 (7.0)	32.4 (6.9)	33.2 (7.1)

Table 2: Demographic characteristics of analytic sample.

Rationale and Description of Exploratory Statistical Analyses

Statistical analyses aimed to explore the following four features associated with the reinforcement schedules on free-living PA. Analyses 2-4 focused only on individuals in the immediate reinforcement group, as only they were exposed to reinforcement schedules.

Analysis 1: Effects of Reinforcement on Physical Activity Relative to Baseline

To assess MVPA during the baseline phase, the cumulative daily bout-minutes were averaged over all participants for each of their last 10 baseline observations (See Fig. 1). Cumulative bout-minutes were

chosen in an effort to mirror the approaches of traditional reinforcement schedule experiments and averages were used to allow us to examine gross effects that would be most useful for informing future trials. The analysis was restricted to each participant’s final 10 baseline days since some participants had more than 10 days (for reasons noted above) and this ensured that the baseline sample assessed all individuals at all time points and was biased towards convergent, non-transient behavior (i.e., less reactivity) expected to be more prevalent at the end of the baseline period.

To determine whether the CFM reinforcement contingency in the immediate reinforcement group implemented at the conclusion of baseline period had an effect on daily MVPA bout-minutes relative to the comparison and the baseline, the average cumulative daily bouts were also calculated for the first 24 intervention observations. A 24-day time interval was selected since this period is the minimum time that individuals in the immediate reinforcement group were required to remain in Stage 1. If a participant met their daily goal on each of these 24 days (e.g., Individual 3 in Fig. 1), they would transition to the next reinforcement stage on Day 25.

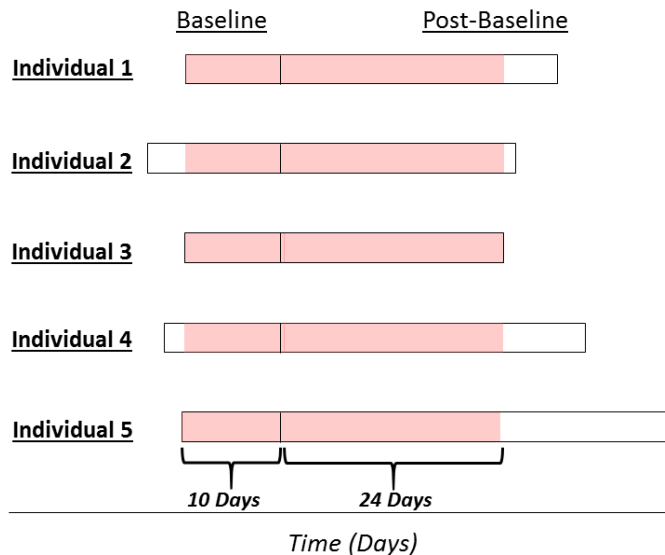


Figure 1: Schematic of the data selected for analysis of baseline versus intervention stages. Total time in the study is illustrated by the black rectangles and the days used within the analysis are highlighted in red. In the baseline period, the last 10 valid days were used and in the intervention period, the first 24 valid days were used.

Analysis 2: Relative Effects of Different Reinforcement Schedules on Physical Activity

A major goal of this ancillary analysis was to assess differences in participants' average cumulative MVPA bouts as a function of their exposure to various reinforcement schedules. Since participants had to reach 24 daily goals to progress out of a stage, the time required to complete this task varies by individual, as shown in Fig. 2a. Therefore, not all participants progressed through all stages and a considerable number ($N=49$) never progressed out of the Stage 1 (See Fig. 2b). Therefore, a simple approach of comparing the mean daily MVPA in each stage (similar to Analysis #1) would average different individuals in each stage. More advanced stages, such as Stages 10 and 11, would be biased towards higher MVPA since only participants who were active enough to progress to later stages would be included in the average.

The above scenario was rectified by stratifying calculations of the mean cumulative MVPA by the *terminal stage* for each participant. Terminal stage was defined as the last stage in which a participant had a minimum of 24 observations, meaning they had an opportunity to advance out of this stage. Individuals who share a terminal stage must, by definition, have also passed through all stages prior to this terminal stage. Therefore, for these participants, the average trajectories of cumulative MVPA bout-minutes for each stage represent the same individuals and are thus comparable. We define groups associated with terminal stage i as TS_i , e.g. TS_3 for individuals with a Stage 3 terminal stage.

The average cumulative MVPA bout-minutes over the last 24 observations in each reinforcement stage was calculated, with results stratified by terminal stage and goal type. We hypothesized that focusing on the end of participants' exposure to each schedule would best reflect the entirety of schedule effects. For each TS /goal type combination, results are presented for participants' average time in the stage (TIS), the average daily goal minutes throughout the stage (GoalMin), and the average number of reinforcements in first 10 days of schedule exposure (Init RE+), which represents initial contact with the

reinforcer. The average trajectory and average number of cumulative bout-minutes (CB) over the last 24 days in the stage were also calculated.

TS1 and TS2 participants were eliminated from these analyses, since they were not exposed to all three types of reinforcement schedule (CRF-FM, CRF-VM, VR-VM). Results were stratified into two subgroups, the first of which consisted of TS3 through TS6 participants. These individuals did not advance far enough through the series of stages to fully encounter each of the six distinct reinforcement schedules; as a result, we call this the *Small Intervention Effect* subgroup. To ensure that the mean MVPA in each stage was calculated over the same individuals, only Stage 1 through 3 averages were calculated for the Small Intervention Effect subgroup. TS7 through TS15 participants comprised the second, *Large Intervention Effect* subgroup. These individuals fully encountered each of the six distinct schedules since the schedules in Stage 7 and beyond had already been used in a previous stage. For this group, MVPA averages were only calculated for Stages 1 through 6 to ensure that the same individuals were being represented.

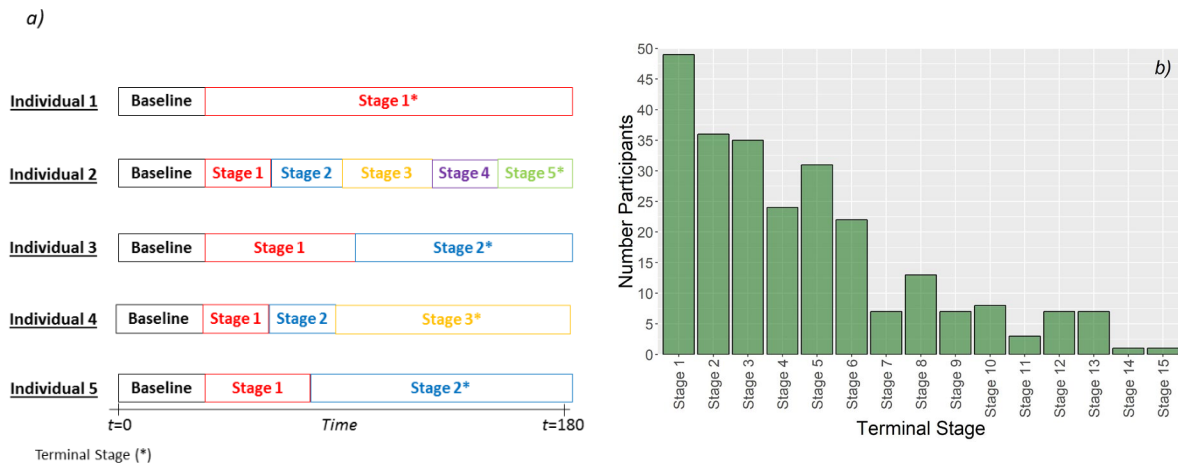


Figure 2: Summary of participants' terminal stages, defined as the last stage in which they spent a minimum of 24 valid days. Panel a) illustrates of schematic of individuals' progression in the intervention showing that the time and total number of stages exposed to varies by participant. Panel b) shows the total number of the 251 immediate reinforcement participants in each possible terminal stage.

Analysis 3: Within-Schedule Temporal Effects

Based on operant theory, it was hypothesized that the novelty of variability within the VM and VR schedules would also maintain participant engagement with the intervention. To investigate whether this occurred, the average cumulative MVPA bout-minutes for the first 7 observations and the final 7 observations within a given schedule were calculated and compared, i.e. $\mu_{last\ 7\ days}^i - \mu_{first\ 7\ days}^i$, where μ_j^i is the average number of MVPA accumulated in Stage i over time period j . All analyses were stratified by goal type and small versus large intervention effect groups.

Analysis 4: Effects of Stage Transition

Because meeting 24 goals was required to advance out of a stage, later stages were, by definition, associated with a longer time in the intervention. Therefore, when examining the effects of the reinforcement schedules in each phase, it can be challenging to distinguish schedule effects from temporal effects. One strategy to isolate schedule effects is to examine the transitions between stages by comparing MVPA patterns at the tail end of one stage versus the beginning of the subsequent stage. This strategy was operationalized by calculating the average MVPA in the first seven days of a stage relative to the last seven days of the previous stage (i.e., $\mu_{first\ 7\ days}^{i+1} - \mu_{last\ 7\ days}^i$) for transitions from Stage1 → Stage 2, Stage 2 → Stage 3, and so forth. This analysis was stratified by goal type and by Small versus Large Intervention effect subgroups, as was done for Analyses 2 and 3.

Results

Analysis 1: Effects of Reinforcement on Physical Activity Relative to Baseline

When examining cumulative daily bout-minutes, consistent activity is represented by a straight line, where the slope represents the rate at which MVPA minutes were accumulated. Therefore, changes between the baseline and reinforcement periods, if present, manifest as differences in slopes over these periods. Fig. 3 indicates that, on average, participants enrolled in both the immediate reinforcement group and comparison groups increased their daily MVPA bout-minutes once the intervention began. After less than five days, participants in the immediate reinforcement group (i.e. Stage 1 [CRF]) were

engaging in more MVPA, on average, than the comparison group. This trend continued throughout the remainder of the intervention period. By the end of the 24th day, the immediate reinforcement group accrued, on average, 116.0 minutes (i.e., 4.8 minutes per day [mpd]) more MVPA than the comparison group.

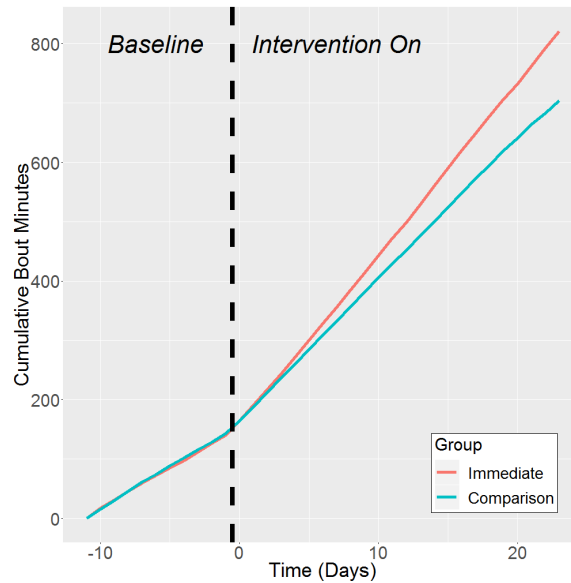


Figure 3: Comparison of cumulative bout-minutes of MVPA in the last 10 days of the baseline versus the first 24 days in the Intervention On period, stratified by financial reinforcement condition.

For the remaining analyses, results are detailed separately for individuals in the Small and Large Intervention Effect subgroups.

Analysis 2: Relative Effects of Different Reinforcement Schedules on Physical Activity

Small Intervention Effect Subgroup (TS3-TS6): Figure 4 illustrates average trajectories and summary statistics for the differential effect of reinforcement schedules/goal types on the amount the cumulative MVPA emitted over the last 24 observations in various stages for individuals in the Small Intervention Effect subgroup (TS3-TS6). Within each goal type, the stage averages were calculated over the same individuals. The static goal group produced more cumulative MVPA than the adaptive goal group (See Figure 4, CB statistic), most likely due to the initial higher goals in the static goal group. For instance, in

Stage 2, the average goal was nearly eight minutes greater in the static versus the adaptive goal group, which led to a difference of 196 minutes of MVPA (8.2 mpd).

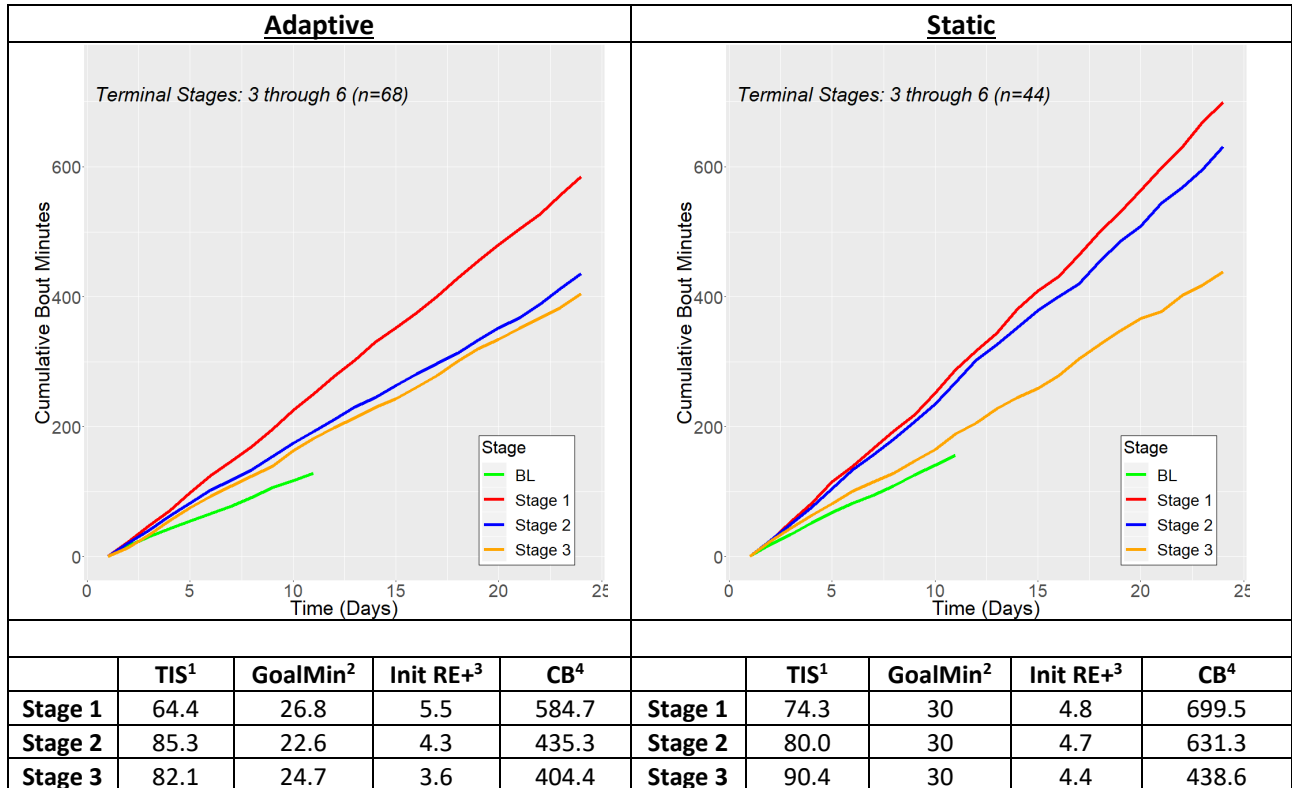


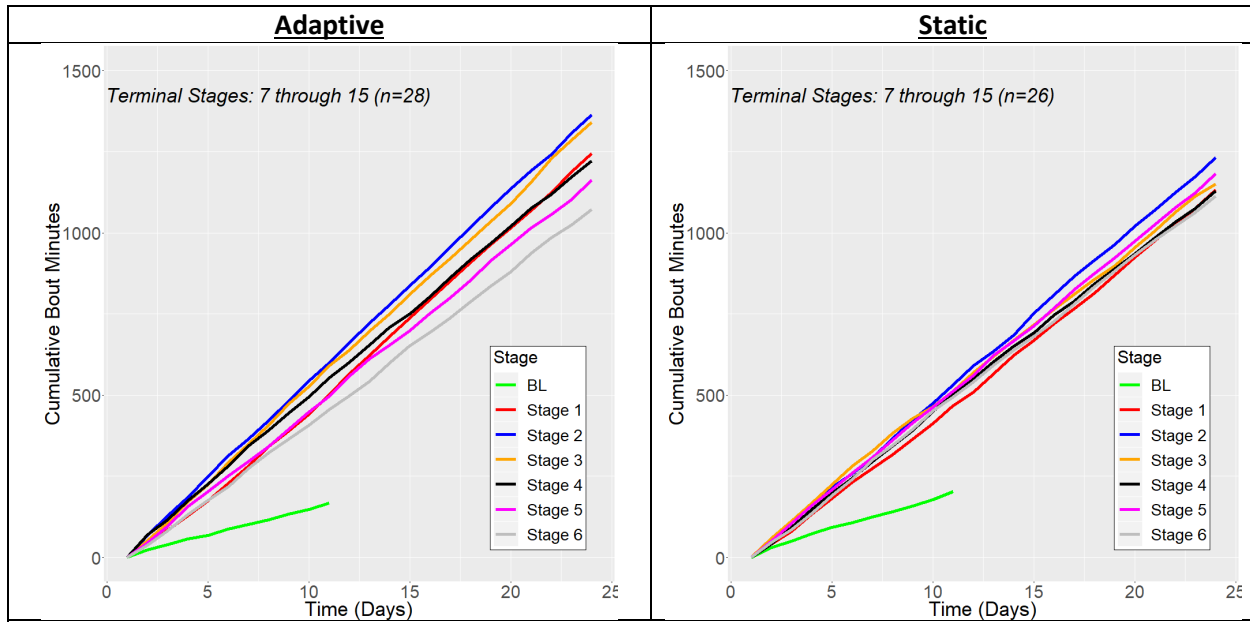
Figure 4: Summary of the differences in MVPA emitted as a function of reinforcement stage for Small Intervention Effect subgroup. The lines in each figure represent average cumulative bout-minutes of MVPA over the last 24 valid days in a stage.

1. TIS: Average time participant spent in stage (days).
2. GoalMin: Average goal for stage (min).
3. Init RE+: Initial contact with the reinforcer, measure by average number of goals met in first 10 days in stage.
4. Average total cumulative bouts in a stage (min).

For the adaptive goal group, Stage 1 clearly produced the most MVPA, but this was not the case for the static group, where Stage 1 and Stage 2 were much more similar. The average performance in the Stage 3, the first VR stage encountered, was comparable for both goal types, although relative to Stages 1 and 2, Stage 3 represented more dramatic reduction in MVPA for the static goal group. Interestingly,

the reduction in MVPA at the end of Stage 3 in the static group occurred despite similar values for initial contact with the reinforcer in Stages 1 and 2.

Large Intervention Effect Subgroup (TS7-TS15): The pattern among stages was qualitatively different for individuals in the Large Intervention Effect subgroup. As shown in Figure 5, adaptive goals produced more MVPA than static goals, with the best stage (Stage 2) producing 131.5 more bout-minutes (5.5 mpd) of MVPA in the adaptive goal group. As was the case for the Small Intervention Effect subgroup, the difference in performance between the adaptive and static groups was likely due to the goal thresholds. For instance, in Stage 2, the average adaptive group goal was of 50.2 min compared to 30 mins in the static group. Despite the larger goals for the adaptive goal group in Stages 1 and 2, initial contact with the reinforcement was greater than for the static goal group. Regardless of goal type, the largest accumulation of MVPA bouts was in Stage 2, during which the first VM reinforcement schedule was presented. For instance, for the adaptive group, 119.4 more minutes (5.0 mpd) of MVPA was produced in Stage 2 compared to Stage 1. This effect was slightly attenuated in the static goal group, where Stage 2 produced 98.6 (4.1 mpd) more MVPA than Stage 1. For the static goal groups, the MVPA accumulation achieved in Stage 1 served as a floor for subsequent reinforcement stages in that MVPA never decreased below Stage 1 levels. This was not the case for the adaptive goal group where Stage 5 and 6 MVPA averages were lower than in Stage 1. In the static group, the initial contact with the reinforcer was relatively consistent over all stages while in the adaptive group, the reinforcer was encountered less frequently as participants progressed through the stages.



	TIS ¹	GoalMin ²	Init RE+ ³	CB ⁴		TIS ¹	GoalMin ²	Init RE+ ³	CB ⁴
Stage 1	31.1	38.7	8.8	1244.0	Stage 1	37.1	30	6.8	1133.3
Stage 2	35.5	50.2	8.0	1363.4	Stage 2	35.0	30	7.7	1231.9
Stage 3	37.3	48.9	7.4	1340.4	Stage 3	35.8	30	7.6	1150.3
Stage 4	40.7	47.8	7.1	1222.0	Stage 4	36.0	30	7.9	1128.9
Stage 5	41	47.9	6.9	1162.8	Stage 5	36.6	30	6.9	1181.9
Stage 6	42.6	44.3	6.9	1072.8	Stage 6	38.0	30	7.7	1112.6

Figure 5: Summary of the differences in MVPA emitted as a function of reinforcement stage for Large Intervention Effect subgroup. The lines in each figure represent average cumulative bout-minutes of MVPA over the last 24 valid days in a stage.

1. TIS: Average time participant spent in stage (days).
2. GoalMin: Average goal for stage (min).
3. Init RE+: Initial contact with the reinforcer, measure by average number of goals met in first 10 days in stage. **Change to match reviewer**
4. Average total cumulative bouts in a stage (min).

Analysis 3: Within-Schedule Temporal Effects

Small Intervention Effect Subgroup (TS3-TS6): The left column of Table 3 illustrates within-stage (i.e., within-schedule) temporal effects for the Small Intervention Effect subgroup. The within-Stage 1 effects were positive for both goal types, indicating more MVPA at the end of Stage 1 than the beginning. This effect was larger in the adaptive goal group. The negative values for Stages 2 and 3 indicate less MVPA at the end of these stages relative to the beginning, which may represent the onset of temporal effects associated with study fatigue. In Stage 3, the first VR-VM reinforcement schedule, the deceleration in MVPA accumulation was much larger for static versus adaptive goals (-11.5 versus -2.9 mpd).

Large Intervention Effect Subgroup (TS7-TS15): The right column in Table 3 illustrates within-stage (i.e., within-schedule) temporal effects for the Large Intervention Effect subgroup. The adaptive goal group generated 85.5 more minutes of MVPA (12.2 mpd) in the last seven days in Stage 1 compared to the first seven days. This effect was attenuated in the static group, who produced 37.5 more minutes (5.4 mpd) of MVPA in the last seven days of Stage 1. Temporal effects in Stage 2 were relatively small compared to those in Stage 1. The VR schedules in Stages 3 through 6 generally resulted in a decelerating trend with less MVPA at the end of stages compared to the beginning. This trend began in Stage 4 for the adaptive goal group and Stage 3 for the static goal group.

	<i>Small Intervention Effect (TS3-TS6)</i>	<i>Large Intervention Effect (TS7-TS15)</i>
	Δ minutes (Δ minutes per day)	Δ minutes (Δ minutes per day)
<i>Adaptive Goals</i>		
<i>Stage 1</i>	20.1 (2.9)	85.5 (12.2)
<i>Stage 2</i>	-34.3 (-4.9)	-14.2 (-2.0)
<i>Stage 3</i>	-20.6 (-2.9)	2.6 (0.37)
<i>Stage 4</i>	-	-34.4 (-4.9)
<i>Stage 5</i>	-	-23.9 (3.4)
<i>Stage 6</i>	-	-50.4 (-7.2)
<i>Static Goals</i>		
<i>Stage 1</i>	5.0 (0.71)	37.5 (5.4)
<i>Stage 2</i>	-13.5 (-1.9)	4.4 (0.63)
<i>Stage 3</i>	-80.8 (-11.5)	-37.8 (-5.4)
<i>Stage 4</i>	-	-21.5 (-3.1)
<i>Stage 5</i>	-	5.0 (0.71)
<i>Stage 6</i>	-	-30.8 (-4.4)

Table 3: Summary of within stage temporal effects. The numbers in this table represent the average change in cumulative bout-minutes of MVPA over the last 7 days in a given stage compared to the first 7 days in this stage. Minutes per day (mpd) representations of these changes are provided in parentheses after each value. Negative numbers indicate that fewer minutes of MVPA were recorded in the last 7 days of a stage versus the first seven days.

Analysis 4: Effects of Stage Transition

Small Intervention Effect Subgroup (TS3-TS6): The left column of Table 4 summarizes the immediate effects on MVPA when transitioning from Stage 1 to Stage 2 and from Stage 2 to Stage 3 for the Small

Intervention Effect subgroup. Larger transition effects were seen for the static goal group, although these effects were small with differences of around 2 mpd of MVPA when transitioning into either Stage 2 or Stage 3.

Large Intervention Effect Subgroup (TS7-TS15): The right column of Table 4 summarizes the immediate effects on MVPA when transitioning from Stages 1 through 5 to the next stage for the Large Intervention Effect subgroup. The transition from Stage 1 (CRF) to Stage 2 (CRF-VM) produced 24.7 more minutes of MVPA (3.5 mpd) for the adaptive goal group, but 6.8 fewer minutes of MVPA for the static goal group. With the exception of the transition from Stage 3 to Stage 4 for the adaptive goal group, all other stage transitions were associated with an immediate increase in MVPA, with larger effects for the adaptive goal group.

For both groups, modifying these analyses by examining different, but comparable, subgroupings (for instance TS2-TS6 rather than TS3-TS6), varying time scales for initial reinforcement contact (e.g., 5 days or 14 days) for Analysis 2, and time intervals of 5, 10, and 14 days at the beginning/end of each phase instead of the 7-day intervals used for analyses 3 and 4 yielded qualitatively similar results. For the interested reader, Appendix 1 contains replications of Figs. 4 and 5 with individual MVPA trajectories shown.

	<i>Small Intervention Effect (TS3-TS6)</i>	<i>Large Intervention Effect (TS7-TS15)</i>
	Δ minutes (Δ minutes per day)	Δ minutes (Δ minutes per day)
<i>Adaptive Goals</i>		
Stage 1 → Stage 2	0.4 (0.06)	24.7 (3.5)
Stage 2 → Stage 3	2.2 (0.31)	13.9 (2.0)
Stage 3 → Stage 4	-	-23.9 (-3.4)
Stage 4 → Stage 5	-	28.5 (4.1)
Stage 5 → Stage 6	-	26.4 (3.8)
<i>Static Goals</i>		
Stage 1 → Stage 2	-9.8 (-1.4)	-6.8 (-1.0)
Stage 2 → Stage 3	16.0 (2.3)	15.0 (2.1)
Stage 3 → Stage 4	-	14.3 (2.0)

Stage 4 →Stage 5	-	16.4 (2.3)
Stage 5 →Stage 6	-	1.4 (0.2)

Table 4: Summary of the transitions into Stages 2,3, and 4 as measured by the change in cumulative bout-minutes of MVPA over the first 7 days of the later stage compared to the last 7 days of the previous stage. Minutes per day (mpd) representations of these changes are provided in parentheses after each value. Negative numbers indicate that fewer minutes of MVPA were recorded in the later stage than the preceding stage.

Discussion

The results from Analysis 1 indicate that the immediate reinforcement group of the WalkIT intervention was successful in increasing MVPA. As expected, immediate, PA-contingent reinforcement produced more MVPA than comparison, non-PA-contingent reinforcement.

Two different subgroups were analyzed based on their overall success in the intervention in terms of the number of goals met. The Small Intervention Effect subgroup (TS3-TS6) was not able to fully progress through all six distinct reinforcement stages in the study while the Large Intervention Effect subgroup (TS7-TS15) did. The results for these two groups were qualitatively different and are detailed and contrasted below.

Small Intervention Effect Subgroup

Stage 1 (CRF-FM) produced the most MVPA for the small intervention effect subgroup, with this effect particularly noticeable for the adaptive goal group, where Stage 2 and 3 averages were both much smaller. Both the adaptive and static goal groups were characterized by large decreases in MVPA within Stages 2 and 3 and small immediate effects when transitioning from one stage to another. These outcomes suggest that individuals for whom the reinforcement intervention had a small effect steadily engaged in less MVPA over time. Static goals were more effective than adaptive goals in delaying this trend.

Large Intervention Effect Subgroup

Stage 2 (CRF-VM) was most successful in producing MVPA for these participants, particularly for the adaptive intervention group where the transition from Stage 1 to Stage 2 immediately yielded an increase in MVPA. This schedule effect may be due the conspicuous change from the CRF-FM schedule in Stage 1 to CRF-VM schedule in Stage 2, which was likely more noticeable than the transitions into the relatively rich VR schedules in subsequent stages. Because the adaptive group had goals that were easier to meet, they likely encountered this novelty more frequently, which may explain the differential effect for these participants. It is notable that the increased level of MVPA was maintained relatively well throughout Stage 2, with little change in MVPA for the first seven days versus last seven days. Static goals were better at reducing MVPA decreases after Stage 2 than were adaptive goals. However, because adaptive goals produced more MVPA initially, physical activity levels were similar for both goal types by the time Stages 5 and 6 were encountered. Stage 1 (CRF-FM) generated some of the lowest levels of MVPA for this subgroup and a floor effect was seen for the static goal participants whereby MVPA levels in all other stages did not decrease below the levels observed in Stage 1. This suggests that the financial reinforcer maintained its potency, even after transitioning to a VR schedule.

The discrepancy between the effects of static versus adaptive goals for the small versus large intervention effect subgroups is likely due to the magnitude of the goals that are presented to participants. Participants in the Small Intervention Effect subgroup did not produce much MVPA and, as a result, do not meet many goals. These participants responded better to a consistent, challenging goal relative to the more lenient goals presented in the adaptive group (Figure 4). It is plausible that an anchoring effect (i.e., evocative effect on MVPA in the presence of an S^D) was present, whereby 30 minutes of MVPA became the acceptable standard that evoked MVPA even when this goal was not met. On the other hand, individuals in the Large Intervention Effect subgroup had much more success in reaching their daily goals, and they appear to have benefited from the fact that adaptive goals could increase beyond 30 minutes of MVPA. These results indicate that future interventions should focus on

identifying small versus large intervention effects as soon as possible so that the optimal goal type can be assigned and/or more potent contingencies can be established. Based on the relative performance of Stage 1 versus Stage 2 for Small versus Large Intervention effects subgroups, sequentially exposing participants to a CRF-FM and then a CRF-VM schedule might serve as a sufficient diagnostic for individualizing interventions. This process also lends itself to the sequential multiple assignment randomized trial (SMART) approach, where performance under intervention arms is assessed for each individual at pre-defined critical points and, where appropriate, intervention strategies for select participants are changed (Collins, Murphy, & Strecher, 2007).

In nearly every case, transitioning into a new VR-VM schedule appeared to increase MVPA, especially for the adaptive goal group. This observation is consistent with operant theory, which predicts that the frequency of a behavior should increase as a schedule is thinned (i.e., as the frequency of behavior required for reinforcement increases). The increased level of MVPA in VR stages was not maintained though, as for both goal types a smaller amount of MVPA was produced over the last 7 days relative to the first 7 days for nearly every VR stage. As exposure to various schedules is inextricably linked to time, it is difficult to discriminate response strain under thinning schedules from temporal effects. Future studies should explicitly test temporal effects by comparing individuals who are kept on a given schedule for an extended period of time versus other individuals who transition to a new schedule.

There were several limitations to the study. First, the nature of the trial design resulted in a form of survival bias, where individuals for whom the intervention was not effective remained in the early stages while others progressed. Ideally, to compare two reinforcement schedules, a random sample of individuals would be exposed to each and then compared. This ideal was not possible for WalkIT Arizona, so we chose to stratify results by terminal stage in order to make sensical comparisons. Second, practical data considerations would not allow accelerometers to continually transmit data to our servers, so device syncing via mobile phone was added as a second component of targeted behavior, which had the potential

to confound pure PA outcomes. Third, VR schedules were implemented via an *a priori* reinforcement protocol that is not perfectly analogous to a standard VR probability approach, which might affect the comparability of our outcomes with previous results. This adjustment was necessary to budget for the year-long study, standardize reinforcement magnitude over study conditions, and appease institutional review board disclosure requirements. Last, a myriad of factors that affect physical activity such as weather, neighborhood walkability, and external events (e.g., injury or vacation) were not included in this analysis.

For much of its history, the intensive observations and manipulation of stimuli required for operant studies have typically forced them to be conducted in laboratory or tightly controlled clinical settings. Even when performed in a natural environment, only the simplest operant features could be explored. However, the newfound ubiquity of mobile and streaming technology presently allows behavior to be observed and environments to be manipulated on a larger scale and with much more precision than in the past. As a result, it is now becoming possible to more faithfully incorporate more subtle operant principles into behavioral interventions within real-world settings, as demonstrated by the WalkIT Arizona study. The most prominent of WalkIT Arizona's operant features was the presentation of positive reinforcement upon participants meeting daily MVPA goals. These goals can be considered a discriminative stimulus that signals the magnitude of MVPA required for reinforcement. To further increase the power of the reinforcer, the time interval between meeting a behavior and receiving reinforcement was kept small by only reinforcing meeting an MVPA goal if the accelerometer was synced by noon the next day. Additionally, all participants regularly received antecedent prompts in the form of text messages that attempted to tap into existing stimulus control (e.g., "Don't like exercising alone? Invite a friend or family member!") and rule-governed behaviors (e.g., "Remember, exercise is heart friendly!"). Taken together, these features increased the study's *theoretical fidelity*, defined as the degree to which a

design adheres to established theory, which has been shown to improve intervention outcomes (Rovniak, Hovell, Wojcik, Winett, & Martinez-Donate, 2005).

WalkIT Arizona represents an example of the ways in which mobile technology enables behavioral interventions to be refined. This technology increases the frequency with which participants can interact with intervention components, allowing methodologies to be tailored in near real-time to participants' behavior and context, and produces a large volume of objective observations that reduces the reliance on unobservable mental constructs in favor of operant considerations. Overall, such developments facilitate a more nuanced approach to promoting healthy behavior that we hope will be more effective than traditional techniques.

Reference List

- Adams, J., Giles, E. L., McColl, E., & Sniehotta, F. F. (2014). Carrots, sticks and health behaviours: a framework for documenting the complexity of financial incentive interventions to change health behaviours. *Health Psychology Review*, 8(3), 286–295. <https://doi.org/10.1080/17437199.2013.848410>
- Adams, M. A. (2009). *A Pedometer-based intervention to increase physical activity: applying frequent, adaptive goals and a percentile schedule of reinforcement*. University of California San Diego.
- Adams, M. A., Hurley, J. C., Todd, M., Bhuiyan, N., Jarrett, C. L., Tucker, W. J., ... Angadi, S. S. (2017). Adaptive goal setting and financial incentives: a 2 × 2 factorial randomized controlled trial to increase adults' physical activity. *BMC Public Health*, 17(1), 286. <https://doi.org/10.1186/s12889-017-4197-8>
- Adams, M. A., Sallis, J. F., Norman, G. J., Hovell, M. F., Hekler, E. B., & Perata, E. (2013). An Adaptive Physical Activity Intervention for Overweight Adults: A Randomized Controlled Trial. *PLoS ONE*, 8(12), e82901. <https://doi.org/10.1371/journal.pone.0082901>
- Adams, M., Hurley, J., Phillips, C., Todd, M., Berardi, V., Hovell, M., & Hooker, S. (2019). Rationale, Design, and Baseline Characteristics of WalkIT Arizona: A Factorial Randomized Trial Testing Adaptive Goals and Financial Reinforcement to Increase Walking across Higher and Lower Walkable Neighborhoods. *Contemporary Clinical Trials*, (In Press).
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman/Times Books/Holt.
- Barte, J. C. M., & Wendel-Vos, G. C. W. (2017). A Systematic Review of Financial Incentives for Physical Activity: The Effects on Physical Activity and Related Outcomes. *Behavioral Medicine*, 43(2), 79–90. <https://doi.org/10.1080/08964289.2015.1074880>
- Berlyne, D. E., Koenig, I. D., & Hirota, T. (1966). Novelty, arousal, and the reinforcement of diversive exploration in the rat. *Journal of Comparative and Physiological Psychology*, 62(2), 222–226.

<https://doi.org/10.1037/h0023681>

- Biglan, A. (2003). Selection by Consequences: One Unifying Principle for a Transdisciplinary Science of Prevention. *Prevention Science, 4*(4), 213–232. <https://doi.org/10.1023/A:1026064014562>
- Brownson, R. C., Boehmer, T. K., & Luke, D. A. (2005). Declining Rates of Physical Activity in the United States: What Are the Contributors? *Annual Review of Public Health, 26*(1), 421–443. <https://doi.org/10.1146/annurev.publhealth.26.021304.144437>
- Chung, S. H., & Herrnstein, R. J. (1967). Choice and delay of reinforcement¹. *Journal of the Experimental Analysis of Behavior, 10*(1), 67–74. <https://doi.org/10.1901/jeab.1967.10-67>
- Collins, L. M., Murphy, S. A., & Strecher, V. (2007). The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART): New Methods for More Potent eHealth Interventions. *American Journal of Preventive Medicine, 32*(5), S112–S118. <https://doi.org/10.1016/J.AMEPRE.2007.01.022>
- Conn, V. S., Hafdahl, A. R., & Mehr, D. R. (2011). Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. *American Journal of Public Health, 101*(4), 751–758. <https://doi.org/10.2105/AJPH.2010.194381>
- Donaldson, J., & Normand, M. (2009). Using goal setting, self-monitoring, and feedback to increase calorie expenditure in obese adults. *Behavioral Interventions: Theory & Practice in Residential & Community-Based Clinical Programs, 24*(2), 78–83. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/bin.277?casa_token=e2oP8GEQxxwAAAAA:qeIna kbH2zs5MwglzgSK8QSRdxcOEGRdr7dpHyvK-Gwg0wYpm49rwpE6Mu3r3szeURLmsAcveQivW730
- Epstein, L. H., Smith, J. A., Vara, L. S., & Rodefer, J. S. (1991). Behavioral economic analysis of activity choice in obese children. *Health Psychology, 10*(5), 311–316. <https://doi.org/10.1037/0278-6133.10.5.311>
- Epstein, L. H., Valoski, A. M., Vara, L. S., McCurley, J., Wisniewski, L., Kalarchian, M. A., ... Shrager, L. R. (1995). Effects of decreasing sedentary behavior and increasing activity on weight change in obese children. *Health Psychology, 14*(2), 109–115. <https://doi.org/10.1037/0278-6133.14.2.109>
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. East Norwalk: Appleton-Century-Crofts. <https://doi.org/10.1037/10627-000>
- Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise, 30*(5), 777–781. <https://doi.org/10.1097/00005768-199805000-00021>
- Galbicka, G. (1994). Shaping in the 21st century: Moving percentile schedules into applied settings. *Journal of Applied Behavior Analysis, 27*(4), 1297861. <https://doi.org/10.1901/jaba.1994.27-739>
- Giles, E. L., Becker, F., Ternent, L., Sniehotta, F. F., McColl, E., & Adams, J. (2016). Acceptability of Financial Incentives for Health Behaviours: A Discrete Choice Experiment. *PLOS ONE, 11*(6), e0157403. <https://doi.org/10.1371/journal.pone.0157403>
- Giles, E. L., Robalino, S., McColl, E., Sniehotta, F. F., & Adams, J. (2014). The Effectiveness of Financial Incentives for Health Behaviour Change: Systematic Review and Meta-Analysis. *PLoS ONE, 9*(3), e90347. <https://doi.org/10.1371/journal.pone.0090347>

- Halpern, S. D., French, B., Small, D. S., Saulsgiver, K., Harhay, M. O., Audrain-McGovern, J., ... Volpp, K. G. (2015). Randomized Trial of Four Financial-Incentive Programs for Smoking Cessation. *New England Journal of Medicine*, 372(22), 2108–2117. <https://doi.org/10.1056/NEJMoa1414293>
- Hovell, M. F., Wahlgren, D. R., Adams, M. A., & Wahlgren, D. . (2009). The logical and empirical basis for the behavioral ecological model. In R. DiClemente, R. Crosby, & M. Kegler (Eds.), *Emerging Theories in Health Behavior and Research* (2nd ed., pp. 347–385). San Francisco, CA: Jossey-Bass.
- Irons, J. G., Pope, D. A., Pierce, A. E., Van Patten, R. A., & Jarvis, B. P. (2013). Contingency management to induce exercise among college students. *Behaviour Change*, 30(2), 84–95. <https://doi.org/10.1017/bec.2013.8>
- Johnston, M. (2016). What more can we learn from early learning theory? The contemporary relevance for behaviour change interventions. *British Journal of Health Psychology*, 21(1).
- Kruger, J., Ham, S., & Kohl III, H. (2005). Trends in leisure-time physical inactivity by age, sex, and race/ethnicity-United States, 1994-2004. *MMWR MorbMortalWklyRep*, 54(39), 991–994.
- Lerman, D. C., Iwata, B. A., Shore, B. A., & Kahng, S. (1996). Responding Maintained by Intermittent Reinforcement: Implications for the Use of Extinction with Problem Behavior in Clinical Settings. *Journal of Applied Behavior Analysis*, 29(2), 153–171.
- McPherson, S., Burduli, E., Smith, C., Herron, J., Oluwoye, O., Hirschak, K., ... Roll, J. (2018). A review of contingency management for the treatment of substance-use disorders: adaptation for underserved populations, use of experimental technologies, and personalized optimization strategies. *Substance Abuse and Rehabilitation, Volume 9*, 43–57. <https://doi.org/10.2147/sar.s138439>
- Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., ... Wood, C. E. (2013). The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. *Annals of Behavioral Medicine*, 46(1), 81–95. <https://doi.org/10.1007/s12160-013-9486-6>
- Miltenberger, R. (2011). *Behavior modification: Principles and procedures*. Boston, MA: Cengage Learning. Retrieved from <https://books.google.com/books?hl=en&lr=&id=je0IAAAQBAJ&oi=fnd&pg=PR4&dq=Behavior+M odification:+Principles+and+Procedures&ots=NAvhQMQIzs&sig=CC8BTL4cKhrBiv4Ly0oLv5mT4vE>
- Mitchell, M. S., Goodman, J. M., Alter, D. A., John, L. K., Oh, P. I., Pakosh, M. T., & Faulkner, G. E. (2013). Financial Incentives for Exercise Adherence in Adults: Systematic Review and Meta-Analysis. *American Journal of Preventive Medicine*, 45(5), 658–667. <https://doi.org/10.1016/J.AMEPRE.2013.06.017>
- Patel, R. R., Normand, M. P., & Kohn, C. S. (2019). Incentivizing physical activity using token reinforcement with preschool children. *Journal of Applied Behavior Analysis*. <https://doi.org/10.1002/jaba.536>
- Prochaska, J., Redding, C., & Evers, K. E. (2015). The transtheoretical model and stages of change. In K. Glanz, B. Rimer, & K. Viswanath (Eds.), *Health behavior: Theory, research, and practice* (5th ed., pp. 125–148). San Francisco, CA: Jossey-Bass.
- Rachlin, H. (1994). *Behavior and mind: The roots of modern psychology*. New York, NY: Oxford University Press. Retrieved from <https://psycnet.apa.org/record/1993-99016-000>

- Rachlin, H. (2000). *The science of self-control*. Cambridge, MA: Harvard University Press.
- Riley, W. T., Rivera, D. E., Atienza, A. A., Nilsen, W., Allison, S. M., & Mermelstein, R. (2011). Health behavior models in the age of mobile interventions: are our theories up to the task? *Translational Behavioral Medicine, 1*(1), 53–71. <https://doi.org/10.1007/s13142-011-0021-7>
- Rovniak, L. S., Hovell, M. F., Wojcik, J. R., Winett, R. A., & Martinez-Donate, A. P. (2005). Enhancing Theoretical Fidelity: An E-mail-based Walking Program Demonstration. *American Journal of Health Promotion, 20*(2), 85–95. <https://doi.org/10.4278/0890-1171-20.2.85>
- Rowlands, A. V., Gomersall, S. R., Tudor-Locke, C., Bassett, D. R., Kang, M., Frayssé, F., ... Olds, T. S. (2015). Introducing novel approaches for examining the variability of individuals' physical activity. *Journal of Sports Sciences, 33*(5), 457–466. <https://doi.org/10.1080/02640414.2014.951067>
- Skender, S., Ose, J., Chang-Claude, J., Paskow, M., Brühmann, B., Siegel, E. M., ... Ulrich, C. M. (2016, June 16). Accelerometry and physical activity questionnaires - A systematic review. *BMC Public Health*. BioMed Central Ltd. <https://doi.org/10.1186/s12889-016-3172-0>
- Skinner, B. (1953). *Science and human behavior*. New York, NY: Macmillan.
- Troiano, R. P., Berrigan, D., Dodd, K. W., Masse, L. C., Tilert, T., & McDowell, M. (2008). Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise, 40*(1), 181–188. <https://doi.org/10.1249/mss.0b013e31815a51b3>
- U.S. Department of Health and Human Services. (2008). Physical Activity Guidelines Advisory Committee Report, 2008, A1-H14.
- U.S. Department of Health and Human Services. (2018). Physical Activity Guidelines for Americans, 2nd edition.
- Ward, B. W., Clarke, T. C., Nugent, C. N., & Schiller, J. S. (2015). *Early Release of Selected Estimates Based on Data From the 2015 National Health Interview Survey*. Retrieved from <http://www.bobmorrison.org/wp-content/uploads/2017/01/cdc-report-on-uninsured-and-other-population-stats.pdf>
- 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>
- Weinstock, J., Barry, D., & Petry, N. M. (2008). Exercise-related activities are associated with positive outcome in contingency management treatment for substance use disorders. *Addictive Behaviors, 33*(8), 1072–1075. <https://doi.org/10.1016/j.addbeh.2008.03.011>
- Weinstock, S. (1954). Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *Journal of Comparative and Physiological Psychology, 47*(4), 318–322. <https://doi.org/10.1037/h0054129>

Appendix 1: Single Individual Trajectories

The results summarized in Figs. 4 and 5 summarize the *average* MVPA accumulated in each stage for the Small Intervention Effect and Large Intervention Effect subgroups, respectively. This appendix presents several iterations of these figures, each with the individual participant trajectories for a single stage added to it. This will allow trends not seen in the average figures to be observed.

Figure 4: Small Intervention Effect Subgroup

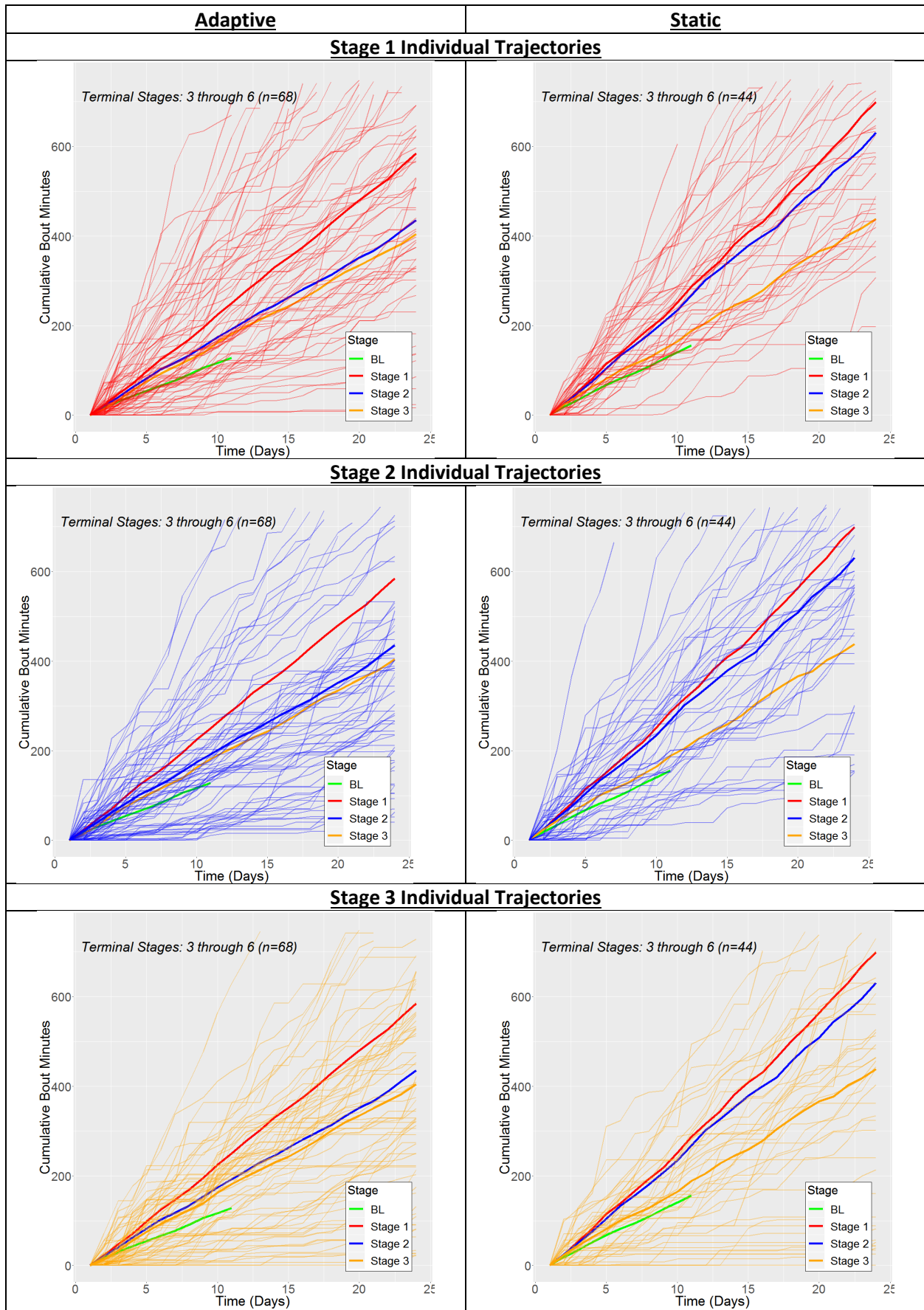


Figure 5: Large Intervention Effect Subgroup

