Tracking valued and avoidant functions with health behaviors: A randomized controlled trial of the acceptance and commitment therapy matrix mobile app

Michael E. Levin, Jennifer Krafft, Seth Seifert & Jason Lillis

Abstract

This randomized controlled trial evaluated the acceptability and additive effects of self-monitoring avoidant and valued functions of behavior, in the context of self-monitoring physical activity and dietary behavior in a mobile app. The self-monitoring approach was based on the Acceptance and Commitment Therapy (ACT) Matrix. A sample of 102 adults interested in improving their diet and physical activity were randomized to a Health Behavior Tracking app (HBT), HBT plus ACT matrix app (HBT+ACT), or waitlist condition. Online self-report assessments were completed at baseline, mid (2 weeks), and post-intervention (4 weeks). Participants reported high usability, but mixed satisfaction with both apps. About half of the prompted app check-ins were completed on average, with 14% never using the ACT app. Participants in the HBT+ACT app condition reported greater self-reported physical activity over time relative to HBT and waitlist, potentially due to protecting against a decrease over time in physical activity observed in the other two conditions. HBT and HBT+ACT conditions both improved self-reported sedentary behavior relative to waitlist. HBT+ACT improved cognitive restraint with eating more than HBT. Neither the HBT or HBT+ACT app improved other health behavior outcome measures or values processes relative to the waitlist. Overall, findings suggest some benefits of the ACT Matrix app for addressing physical activity by tracking valued/avoidant functions, but mixed findings on acceptability, outcomes, and processes of change suggests impact may be relatively limited.

Keywords: Acceptance and Commitment Therapy; mHealth; Obesity; Exercise; Diet

Tracking valued and avoidant functions with health behaviors: A randomized controlled trial of the acceptance and commitment therapy matrix mobile app

Poor diet and physical inactivity are key predictors of physical health problems and early mortality, while healthier diets and regular exercise are known to have myriad benefits to health and well-being (Loef & Walach, 2012; Mozaffarian et al., 2008; Reiner, Niermann, Jekauc, & Woll, 2013; Wang, Li, Chiuve, Hu, & Willett, 2015). Yet, it is difficult to improve diet and physical activity (Forman & Butryn, 2015; Gatewood et al., 2008; Salmon, Crawford, Owen, Bauman, & Sallis, 2003).

Self-monitoring (i.e., observing and recording one's own behavior in an ongoing, structured format) is an established, evidence-based method for improving diet and physical activity (Korotitsch & Nelson-Gray, 1990; Lieffers & Hanning, 2012; Michie, Abraham, Whittington, McAteer, & Gupta, 2009). The ease and availability of self-monitoring diet and physical activity has been substantially improved through mobile apps, which include features such as automated prompts and tools to reduce barriers to self-monitoring such as response effort, forgetting, and knowing how to enter relevant data (Payne, Lister, West, & Bernhardt, 2015; Peng, Kanthawala, Yuan, & Hussain, 2016). However, the efficacy of self-monitoring mobile apps is still limited, with a notable portion of individuals failing to initiate or sustain healthier diet and physical activity (e.g., Carter, Burley, Nykjaer, & Cade, 2013).

One potential avenue for increasing the efficacy of self-monitoring apps is to expand what users monitor. Typically, self-monitoring apps include recording topographical features of behavior such as quantity and type of food consumed and length and intensity of physical activity. Monitoring these behaviors is theorized to improve diet and physical activity by providing baseline information to inform realistic goal-setting, and allowing individuals to

evaluate their progress toward their goals over time, supporting goal-focused behavior (Bandura, 1998).

Recent developments within cognitive behavioral methods highlight an additional aspect of behavior to self-monitor – the function of behavior. Function refers to the antecedents and consequences that influence behavior (i.e., why an action is taken). A variety of topographically distinct behaviors can serve similar functions (e.g., eating highly palatable foods, watching TV, restrictive dieting, avoiding exercise, and so on, can all serve to avoid unwanted emotions), and similar behaviors can serve notably distinct functions (e.g., exercising to improve physical fitness or to escape body shame). Monitoring function can serve to identify and modify maladaptive patterns of diet and physical activity that otherwise may be missed if only tracking topography.

In particular, this study focused on two functional classes of behavior specified by Acceptance and Commitment Therapy (ACT; Hayes, Strosahl, & Wilson, 2011), a modern cognitive behavioral therapy that has been found effective for improving diet, physical activity, and weight loss outcomes in over 20 clinical trials to-date (e.g., Forman et al., 2016; Lillis, Hayes, & Levin, 2011; Lillis & Kendra, 2014; Moffitt & Mohr, 2015; Niemeier, Leahey, Reed, Brown, & Wing, 2012). ACT focuses on decreasing behaviors that are excessively and harmfully focused on changing aversive internal states (i.e. behaviors that function to reduce, escape, or replace unwanted thoughts, feeling, or bodily sensations, also known as experiential avoidance). A variety of maladaptive behaviors related to diet and exercise serve experientially avoidant functions such as overeating or eating high caloric foods in response to aversive internal states (e.g., emotional eating, binge eating, disinhibited eating), sedentary behavior to avoid aversive internal states (e.g., distracting with TV, avoiding discomfort from exercise), and excessive,

ineffective attempts to lose weight quickly in an effort to avoid aversive states like shame (e.g., overly restrictive dieting or fasting; Lillis & Kendra, 2014). Research has consistently found experiential avoidance leads to poor diet and physical inactivity (e.g., Lillis et al., 2011; Litwin, Goldbacher, Cardaciotto, & Gambrel, 2017; Palmeira, Cunha, & Pinto-Gouveia, 2018; Ulmer, Stetson, & Salmon, 2010), and that reducing experiential avoidance through ACT leads to improvements in diet and weight loss (e.g., Forman et al., 2016; Lillis et al., 2011; Niemeier et al., 2012).

In addition to decreasing experiential avoidance, ACT seeks to promote values-based behaviors, which are behaviors under the control of verbally specified reinforcers that are personally meaningful, intrinsically reinforcing qualities of action. Research has consistently found that intrinsic positive motivators can increase healthy diet and physical activity (Teixeira, Carraca, Markland, Silva, & Ryan, 2012; Vestuyf et al., 2012) and mediate treatment effects for ACT on weight loss (e.g., Forman et al., 2016).

ACT typically targets experiential avoidance and values-based behaviors through in-person interventions teaching acceptance, mindfulness, and values-based skills, but this approach is fairly intensive and requires trained providers, which may limit reach and integration into other health behavior change efforts. Recently, a lower intensity, alternative ACT protocol has been developed called the ACT Matrix, which focuses primarily on self-monitoring one's overt behaviors and internal experiences in relation to experiential avoidance and values (Polk & Schoendorff, 2014). Theoretically, practicing such self-monitoring leads to decreased experiential avoidance and increased values-based activities by increasing awareness of the function of one's behavior, identifying experientially avoidant behaviors that are ineffective or inconsistent with personal goals, and noticing opportunities and increasing motivation for

healthy behaviors connected to personal values. Although the ACT Matrix provides a potentially simpler and more focused approach that could be integrated into self-monitoring interventions to augment health behavior change it has received minimal research to-date.

Such a low intensity intervention might best be implemented in a self-guided mobile app format, mirroring the use of mobile apps for delivery of self-monitoring of diet and physical activity (Payne et al., 2015). Although relatively preliminary, there is a growing body of research demonstrating the potential efficacy of ACT when delivered through mobile apps (Linardon, in press; Torous, Levin, Ahern & Oser, 2017). For example, ACT mobile apps have been found efficacious for smoking cessation (Bricker et al., 2014), depression and anxiety symptoms (Levin, Haeger & Cruz, 2019), and health behaviors (Levin, Pierce & Schoendorff, 2017).

Of most relevance to the current study, the ACT Matrix was evaluated as a mobile app in a pilot randomized controlled trial (RCT) with 23 adults interested in improving their diet and physical activity (Levin et al., 2017). Participants were randomized to a waitlist condition or to use the ACT Matrix App for two weeks after a brief in-person orientation on how to use the app. The app was relatively simple, consisting of random prompts three times a day in which users were asked to self-monitor if their current behavior was a toward move (i.e., action moving towards personal values) or away move (i.e., action moving away from aversive internal states). Participants in the ACT Matrix condition who regularly engaged with the app (75%, n = 9) improved more on health behaviors relative to the waitlist (d = 1.04), but this effect was not found in the full intent-to-treat sample (i.e., everyone assigned to use the app). Participants also improved on the frequency of reported toward moves (i.e., valued actions) in the mobile app over time, although not on self-reported valued behavior at post. These results suggested the potential benefits of tracking the function of behaviors as a means of improving diet and physical activity,

but the initial pilot was limited by the short two-week assessment window, small sample size, and lack of an active control group.

Thus, the present study sought to further evaluate the efficacy of the ACT Matrix app on diet and physical activity health behaviors over four weeks with a sample of 102 adults against an active comparison condition. Participants were randomized to a health behavior tracking app focused on daily self-monitoring of diet and physical activity (HBT), the ACT Matrix app which included the same HBT tracking features plus ACT Matrix features focused on monitoring the function of health behaviors (HBT+ACT), or a waitlist condition. The first study prediction was that the revised version of the ACT Matrix app would be highly acceptable to participants as indicated by high rates of program usage and self-reported program satisfaction. The second prediction was that the HBT+ACT condition would produce greater improvements in self-reported diet and physical activity than the HBT or waitlist conditions. The third prediction was that HBT+ACT would produce greater improvements in valued action as the key process of change than the HBT or waitlist conditions. If these predictions were supported it would suggest that tracking the function of health behaviors, in addition to the topography of behaviors, through approaches like the ACT Matrix can serve to improve diet and physical activity.

Methods

Participants

A final sample of 102 adults interested in improving their diet and physical activity were enrolled in the study. Eligibility requirements included being 18 years of age or older, owning an iOS or Android smartphone, and having an interest in making diet and exercise behavior-related changes in their lives. Recruitment was primarily conducted at a mid-sized university in the Mountain West region of the United States (e.g., flyers, online advertisements). Potential

participants were directed to an online screener for various clinical trial opportunities offered through the laboratory.

The sample was 75% female with an average age of 23.51 (*SD* = 8.07, median = 21, range = 18-57). The sample was predominantly non-Hispanic White (88%), with 4% Hispanic White, 3% Hispanic non-White, 2% Asian, 2% Black/African American, and 1% multiracial. In terms of highest education achieved, 65% completed some college (but not yet received a degree), 12% had a Bachelor's degree, 10% completed high school or equivalent degree, 8% completed a technical degree, 4% Master's degree, and 1% doctoral degree. The average self-reported BMI at baseline was 24.94 (*SD* = 4.44, range = 18.07-39.05), with 41% in the overweight or obese category (BMI $\geq$ 25). Based on empirically derived cutoff scores on the *International Physical Activity Questionnaire* (IPAQ; Craig et al., 2003), participants generally reported being highly physically active, with 69% meeting the "high" physical activity criteria, 30% "moderate," and only 1% "low."

Procedures

All study procedures were completed online, primarily through the Qualtrics and LifeData research platforms. After a brief online screening for eligibility criteria (age, owning an iOS or Android smartphone, interest in making diet and exercise behavior-related changes), participants were automatically directed to online informed consent. An online baseline self-report survey then assessed demographics and all outcome and process measures besides program satisfaction. Participants were automatically randomized by Qualtrics after completing the baseline survey to one of three conditions: Health Behavior Tracking App (HBT), HBT plus ACT Matrix App (HBT+ACT), or Waitlist. Block randomization, which consisted of 12 slots with 4 per condition that reset each time the 12th participant was randomized, was used to ensure

a relatively equal sample size over time. Participants assigned to an app condition (HBT or HBT+ACT) were given a self-guided online tutorial and instructed to use the app for the following four weeks. Waitlist participants simply waited four weeks, receiving access to either app after completing the final post assessment. Online surveys were administered at mid (2 weeks after baseline) and post (4 weeks after baseline) assessment time points, which included all outcome and process measures besides program satisfaction, which was only assessed at post. Participants who were college students in relevant psychology courses were given the opportunity to receive course credit through the Sona research platform for completing the online assessment portions of the study. No other incentives were provided for research participation. The study was approved by the authors' institutional review board.

HBT Condition

Prior to installing the HBT app, participants were directed to complete an online tutorial delivered through Qualtrics. We have regularly used Qualtrics to provide self-guided behavioral interventions and app tutorials, as it offers a wealth of features to provide engaging, interactive content (e.g., Levin et al., 2017; Levin et al., 2019). The tutorial was estimated to take 5-10 minutes to complete, covering content including identifying personal health goals for using the app, the benefits of self-monitoring, how to use the app, and defining and clarifying each of the behaviors to be tracked in the app (including what would make a meal healthy or unhealthy, exercise, and sedentary behavior, and what their health goals are to work on each day). After completing the tutorial, participants were provided instructions for how to download the HBT app.

The HBT app was developed and delivered through the LifeData research platform. LifeData provides an effective, easy-to-use platform to develop and deploy prototype native

mobile apps that users can download and access from their mobile phone. The HBT app consisted of only an evening check-in, which participants were expected to complete daily. The evening check in was only available when prompted each evening at 9pm (for up to 24 hours until completed). This evening check-in included a total of nine self-monitoring questions. The first four questions asked participants whether each meal for the day (breakfast, lunch, snacks, dinner) was either healthy or unhealthy as defined by the user based on features such as portion size, calories, and nutritional value. These questions were purposefully broad to minimize the burden of completing daily self-monitoring and maximize flexibility in relation to personalized dietary goals, rather than asking for specific details regarding caloric intake and specific food groups, which require a much higher response effort. The next two questions asked participants if they engaged in "a planned/structured physical activity for at least 20 minutes or more that made you sweat or breathe harder than normal" with yes or no response options. Participants who said yes were then asked how much time they spent in this physical activity from 20 minutes to "more than 2 hours." The next two questions ask participants to monitor their sedentary behavior including the "longest, uninterrupted period of sedentary activity" (from "30 minutes or less" to "more than 3 hours") and the percentage of their day spent doing sedentary activities (from "a small portion of the day [20% or less]" to "almost all day [over 80%]"). The final question asked participants if they met their health goal today with response options including yes, partially, or no. These health goals were initially set during the app tutorial, but participants were instructed to modify their goals as needed throughout the four week period.

Participants received templated email contacts from a research assistant at multiple time points throughout the four weeks to continue using the app. This included standard check-in prompts 2 and 7 days after completing the online orientation to continue using the app as well as

prompting up to 3 times every 3-4 days if a user stopped using the app. Participants were also sent an email summary of their monitoring data at the end of each week to support ongoing monitoring and identifying patterns over time in eating and physical activity data.

HBT+ACT Condition

The ACT Matrix app was also delivered in LifeData and included all of the same components as the HBT condition, including an online tutorial, evening check-ins that include self-monitoring of eating and physical activity, and email check-ins and summaries from a research assistant. In addition to the HBT components, the ACT matrix included additional content and components focused on monitoring valued functions of behavior (i.e., toward moves: behavior that moves towards personal values) and avoidant functions of behavior (i.e., away moves: behavior that moves away from aversive internal experiences).

Prior to downloading the ACT app, participants completed a similar online tutorial as the HBT condition. In addition to what was covered for the HBT app, participants were also oriented to values and avoidant patterns of behavior. This included identifying personal values related to health, behaviors one might take to move towards these values, aversive internal experiences (e.g., thoughts, feelings, sensations), and behaviors one might take to move away from these aversive internal experiences (e.g., avoidance). The tutorial explored these patterns of behavior, how people get stuck in unhelpful away move patterns, and how to notice when one is engaging in a toward versus away move in the moment.

The ACT Matrix app included the same evening check-in as the HBT app, with additional questions focused on monitoring the function of health behaviors as part of the check-in. For each meal, participants were asked to also rate whether their choice was a toward or away move, in addition to whether it was healthy or unhealthy. Similarly, in addition to monitoring

if/how much exercise they engaged in, participants were asked to rate whether their "choices with exercise (or to not exercise)" were more of an away move or a toward move. Finally, the ACT Matrix app included two additional questions at the end regarding "the level of inner obstacles you experienced today in relation to your health behaviors and goals" (sliding scale from 0 "none" to 10 "a lot") and "how much were you able to move toward who/what is important to you with your health, even in the presence of inner obstacles (sliding scale from 1 "never able to move toward" to 10 "always able to move toward").

The ACT Matrix app also included a check-in that users were prompted to complete each morning at 9am. The morning check-in was designed to increase participants' awareness and monitoring of the functions of their health behaviors throughout the day. Although monitoring the function of behavior after it occurs is a key part of the ACT matrix, ideally participants learn to notice toward and away moves when they occur in the moment (or before they occur), which further supports behavior change. The morning check-in included questions to identify health-related values, behaviors that would move towards those values, internal barriers that might arise, and potential away moves. Of note the morning and evening check-ins were not used in our previous ACT Matrix app pilot trial (Levin et al., 2017), but were added based on participant feedback in the pilot that the app could be more tailored to health behavior change efforts as well as findings from another study that adding additional evening check-ins and related tools increased the efficacy of the ACT Matrix app for well-being (Krafft, Potts, Schoendorff, & Levin, 2019).

Finally, participants were prompted randomly three times a day between 10am and 8pm to check in on toward and away moves. These check-ins mirrored the prompts used in our initial ACT Matrix pilot trial (Levin et al., 2017), in which participants were asked "Right now are you

engaged more in an away move (away from what you don't want to think and feel) or a toward move (toward who or what is important). If not sure, just guess," with response options for "toward" or "away." Additional questions were added that were not used in the initial ACT Matrix pilot trial given feedback from participants that the previous app was too simple (Levin et al., 2017), and positive results in a subsequent trial comparing the simpler matrix app to a version with additional questions/features on general well-being (Krafft et al., 2019). Participants were asked to pick a value they are moving towards or to pick an internal barrier they are trying to move away from using a list of examples. Participants were also asked to consider if they had taken any toward moves with their health values since their last check-in, if they had taken any away moves related to their health, and if they had met their health values goal for the day that they set during the morning check-in.

The evening and morning check-ins were only available when prompted (for up to 24 hours until completed). The toward and away move check-in was available to be completed at any time in addition to the random daily prompts to complete a check-in.

Outcome Measures

International Physical Activity Questionnaire (IPAQ; Craig et al., 2003). The IPAQ was used to assess total physical activity and sedentary time. The IPAQ includes a series of questions assessing the frequency and intensity of physical activity in life domains including work, household chores, leisure, and commuting. Respondents are asked to estimate how many days during the last week they engaged in physical activity of varying intensity in these domains (vigorous exercise, moderate exercise, walking) as well as how many minutes they spent on the activity each day. Total physical activity is calculated by weighting the amount of physical activity based on intensity and summing it into a total activity score. In addition to a continuous

score, the IPAQ data can be used to calculate categorical scores for "low," "moderate," and "high" physical activity. Participants are also asked to self-report the amount of sedentary time spent on average over weekdays and weekends, which are summed into a sedentary total score. The IPAQ has demonstrated good psychometric properties (Craig et al., 2003) and acceptable concurrent validity with physical activity logs and the recordings of electronic activity monitors (Craig et al., 2003; Hagströmer, Oja, & Sjöström, 2006). The IPAQ has been criticized at times for inconsistency with objective measures (Lee, Macfarlane, Lam, & Stewart, 2011); however, this is an issue common across self-report measures of physical activity (Sallis & Saelens, 2000). In addition, test-retest reliability of the IPAQ is acceptable (Craig et al., 2003), which has led even researchers critical of the validity of the IPAQ to note that it may be useful in characterizing whether longitudinal change occurred, if not its magnitude (Lee et al., 2011).

Weight Control Strategies Scale (WCSS; Pinto, Fava, Raynor, LaRose, & Wing, 2013). The 30-item WCSS was used to assess key health behaviors related to weight management including dietary choices, self-monitoring strategies, physical activity, and psychological coping. Items are rated on 5-point scales ranging from "Never" to "Always." A total score can be calculated by summing all individual items, with higher scores indicating overall healthier behaviors for managing weight (e.g., healthier dietary choices, greater physical activity, more adaptive coping strategies to manage weight, greater self-monitoring of weight-related behaviors). The WCSS has been found to be a reliable and valid measure and to be sensitive to the impact of treatments for weight loss (Pinto et al., 2013). For example, the WCSS dietary choices subscale improved following behavioral weight loss treatment and predicted both weight loss and total daily caloric consumption (Pinto et al., 2013). The WCSS was also found to

improve in a prior evaluation of the ACT Matrix app (Levin et al., 2017). Internal consistency was good for the WCSS in this sample ($\alpha = 0.90$).

Three Factor Eating Questionnaire-Revised (TFEQ; Karlsson, Persson, Sjöström, & Sullivan, 2000). The TFEQ is an 18-item measure of three types of eating patterns: cognitive restraint (i.e., control over eating), uncontrolled eating (i.e., difficulty regulating eating), and emotional eating (i.e., overeating when experiencing negative mood). Most items are rated on a 4-point scale from "Definitely true" to "Definitely false," but response options vary by item. The TFEQ subscales have demonstrated adequate internal consistency and convergent validity with reported food intake in a general population (de Lauzon et al., 2004). In the present study, reliability was acceptable for cognitive restraint ($\alpha = 0.77$), good for uncontrolled eating ($\alpha = 0.87$), and good for emotional eating ($\alpha = 0.85$).

General Health Questionnaire (GHQ; Goldberg, 1972). The 12-item GHQ was used to assess general psychological distress as a secondary outcome. Although the focus of the ACT Matrix app was on improving health behaviors, it teaches skills that are more broadly applicable to improving well-being and quality of life. GHQ items were rated on a 4-point scale with lower scores indicating less distress. The GHQ has been found to have adequate reliability and validity in past studies (Banks, 1980). Internal consistency was good in the present sample ($\alpha = 0.88$)

Process Measures

Valued Living Questionnaire (VLQ; Wilson, Sandoz, Kitchens, & Roberts., 2010). The 20-item VLQ was used to examine values-consistent action in key life domains, a process predicted to improve through the ACT Matrix app. This measure asks participants to rate how important each of 10 life domains has been over the past week (from 1 "not at all important" to 10 "extremely important") and then how consistent one's actions have been in each of these

domains over the past week (from 1 “not at all consistent” to 10 “extremely consistent”). A total score is calculated by multiplying importance by consistency, indicating to what extent individuals are behaving consistently with domains they identify as important. The VLQ has been found to be a reliable and valid measure in past research (Wilson et al., 2010). Reliability was adequate for both importance ($\alpha = 0.77$) and consistency ($\alpha = 0.71$) in the current study.

Committed Action Questionnaire (CAQ; McCracken, Chilcot, & Norton., 2014). The 8-item CAQ was used to examine committed action (i.e., persisting in actions related to personal values), another key psychological process predicted to improve from the ACT Matrix app. Each item is rated on a 7-point scale from 0 “never true” to 6 “always true.” The CAQ has been found to be reliable and valid in preliminary research (McCracken et al., 2014), and had good internal consistency in this study ($\alpha = 0.88$).

System Usability Scale (SUS; Brooke, 1996). The SUS is a 10-item scale that measures usability, a key facet of program acceptability and satisfaction with regards to how easy and enjoyable a program is to use. The SUS was administered only to participants assigned to one of the active conditions at the post-intervention time point. Each item is rated from 1 “Strongly disagree” to 5 “Strongly agree.” The SUS has empirically derived benchmarks for “Excellent” and “Good” usability (Bangor, Kortum & Miller, 2008). A review of SUS research in 206 studies found strong support for the factor structure and reliability of this measure (Bangor et al., 2008). The SUS has been widely used to evaluate the usability of behavioral health apps, including previous pilot trials of the ACT matrix app (Krafft et al., 2019; Levin et al., 2017), thus allowing for comparisons across studies. Internal consistency was good for the SUS in the current study ($\alpha = 0.87$).

Individual Satisfaction Items. In addition to the SUS, a series of individual program satisfaction items were provided at post-intervention for participants assigned to one of the active conditions. These items were adapted from previous studies evaluating ACT mobile apps, including the ACT matrix app (Krafft et al., 2019; Levin et al., 2017). Items assessed domains of program satisfaction including overall satisfaction, perceived helpfulness, desire to use the app in the future, and usability. Each item was rated on a 6-point forced choice scale (1 *strongly disagree,* 2 *disagree,* 3 *slightly disagree,* 4 *slightly agree,* 5 *agree,* 6 *strongly agree*). A 4 (*slightly agree*) or higher provided a benchmark for indicating a positive response, which although not empirically validated directly, has been used due its face validity to determine program satisfaction in prior studies (e.g., Krafft et al., 2019; Levin et al., 2017*).*

Analytic Plan

Recruitment was conducted for one year (February 2017 to March 2018), with the aim of recruiting at least 60 and up to 120 participants. Although an a prior power analysis was not used to select the achieved sample, a post hoc power analysis indicated the final sample of 70 participants who completed the post-intervention assessment provided adequate power (.80) to detect a small effect size (d = .34) at $p < .05$ for a time by condition test with three time points and three conditions.

Descriptive statistics were examined to determine rates of program usage and self-reported acceptability/satisfaction with the mobile apps. Independent sample *t*-tests compared program usage and satisfaction ratings between the HBT and HBT+ACT conditions.

Mixed model repeated measure (MMRM) analyses tested for differences between conditions over time on each outcome and process measure. Analyses were conducted with the full intent-to-treat sample that completed the baseline assessment and were assigned to study

condition, with MMRM able to model available data even if mid and/or post assessments were missing for some participants. Each model included effects for time (baseline, mid, post) and condition (HBT, HBT+ACT, Waitlist) and time by condition interactions. Significant time by condition interactions were further examined with post hoc tests for within condition improvements and between condition differences at specific time points. If no significant post hoc tests were found when examining a significant time by condition test, then trending ($p < .10$) post hoc effects were interpreted. If any significant differences were found between conditions at baseline, these variables were planned for inclusion as covariates in MMRM analyses.

In addition, a generalized linear mixed-effects model (GLMM) with a logit link was run to test the hypothesis that the rate of toward moves would increase over time within the HBT+ACT condition based on user-entered data in the app. Random intercepts were modeled at the participant level, to account for multiple observations nested within participants. Parameters were estimated with the Laplace approximation of maximum likelihood, which uses all available data. This approach allows for appropriate modeling of binomial data (i.e., participants either moved toward or away).

Results

Preliminary analyses

The distribution of each outcome and process variable was checked relative to a normal distribution. The IPAQ physical activity, IPAQ sitting time, and GHQ mental health variables were each significantly skewed. A logarithmic transformation was employed for IPAQ sitting time and a square root transformation for IPAQ physical activity and GHQ mental health to approximate a normal distribution.

Overall, 75% of the sample completed mid and/or post assessments (62% of the sample completed the mid assessment and 69% completed post). There were no significant differences in missing data rates between conditions (see Figure 1 participant flow diagram for missing data rates per condition).

Baseline differences between conditions were examined through a series of ANOVA and chi square analyses. There were significant differences between conditions at baseline on CAQ committed action, $F(2,99) = 3.96$, $p = .022$, and BMI, $F(2,99) = 4.64$, $p = .012$. At baseline, participants in the HBT condition reported higher committed action than those in the waitlist (*Mdiff* = 3.52, $p = .044$) or HBT+ACT conditions (*Mdiff* = 4.64, $p = .008$). At baseline participants in the HBT condition also reported lower BMI than those in the waitlist (*Mdiff* = -3.18, $p = .003$), but not HBT+ACT conditions (*Mdiff* = -1.88, $p = .075$). Baseline BMI and CAQ variables were thus included as covariates in analyses comparing between group effects over time. There were no other differences between conditions at baseline on outcome or demographic variables.

Program Usage

The online tutorial that participants were required to complete in order to use their assigned app was completed by 97% in the HBT+ACT condition and 85% in the HBT condition. Overall, 86% used the HBT+ACT app at least once and 76% used the HBT app at least once. Out of 28 evening diaries, on average participants in the HBT+ACT condition completed 16.71 (*SD* = 10.12, 69% completing 14 or more). Participants in the HBT condition completed 15.46 evening diaries (*SD* = 11.66, 58% completing 14 or more). There were no differences between conditions on program usage ($p > .05$). Participants in the HBT+ACT condition completed

almost half of the 56 random values prompts (*M* = 26.11, *SD* = 18.29), which were not provided in the HBT condition.

These usage rates are comparable to those found in past mobile app studies. A review of 57 mobile app studies found 0%-58% of participants did not download or ever use the assigned app, with an average of 21% never using the app for those targeting general mental health (Linardon & Fuller-Tyszkiewicz, in press). The number of completed check-ins were also similar to those found in prior ACT matrix app studies, which have ranged from 13 to 18 app interactions a week (Krafft et al., 2019; Levin et al., 2017).

Program satisfaction

SUS ratings were equivalent for the HBT app and HBT+ACT app, $t(43) = 1.78$, $p = .083$; HBT *M* = 82.95, *SD* = 13.23, HBT+ACT *M* = 76.17, *SD* = 12.38. These scores are in the "good" to "excellent" range based on empirically derived cutoff scores for the SUS (Bangor et al., 2008), and within 1 SD of the SUS score for our original ACT matrix app study (*M* = 82.50, *SD* = 10.25; Levin et al., 2017).

Most satisfaction ratings were on the borderline between slight satisfaction (4 = "slightly agree") and slight dissatisfaction (3 = "slightly disagree"): "Overall, I was satisfied with the app" (HBT+ACT *M* = 4.08, *SD* = 1.21; HBT *M* = 4.33, *SD* = 1.39), "The app helped me with my diet and/or exercise goals" (HBT+ACT *M* = 3.42, *SD* = .88; HBT *M* = 3.90, *SD* = 1.64), and "I would use the app again in the future" (HBT+ACT *M* = 3.08, *SD* = 1.41; HBT *M* = 3.24, *SD* = 1.48). However, consistent with higher SUS usability ratings, both apps were rated highly on ease of use ("The app was easy to use;" HBT+ACT *M* = 5.08, *SD* = 1.10; HBT *M* = 5.29, *SD* = 1.10). Almost all satisfaction ratings were equivalent between conditions. Participants in the HBT condition more strongly disagreed with the statement "I wouldn't have been able to use the app

without the orientation I completed on the website" (i.e., they thought the online tutorial was not necessary), $t(43) = 2.06$, $p = .045$; HBT+ACT $M = 3.25$, $SD = 1.45$; HBT $M = 2.33$, $SD = 1.53$.

Satisfaction ratings were within 1 *SD* of ratings given on similar items in the initial ACT matrix app pilot (Levin et al., 2017) for overall satisfaction ($M = 4.18$), perceived helpfulness ($M = 3.55$), and desire to use the app in the future ($M = 3.82$). However, these ratings are lower than those found on similar items with other ACT mobile apps (e.g., Levin et al., 2019) for overall satisfaction ($M = 5.23$), helpfulness ($M = 5.23$), and desire to use the app ($M = 5.15$).

MMRM outcome and process analyses

A series of MMRM analyses tested for time by condition effects between the three conditions (HBT+ACT, HBT, Waitlist) and time points (pre, mid, post), while controlling for baseline CAQ committed action and BMI as covariates (see Table 1 for descriptive statistics by condition and time point).

A significant time by condition interaction was found for IPAQ physical activity, $F(4,67.29) = 2.81$, $p < .05$, which appeared due to participants in the HBT+ACT condition reporting greater physical activity over time than HBT or Waitlist. There were no significant post hoc comparisons within or between conditions. However, there were trends for the HBT+ACT condition having greater IPAQ physical activity at post than the HBT, $t(59.93) = 1.74$, $p = .088$, $d = .52$, or Waitlist conditions, $t(54.24) = 1.68$, $p = .098$, $d = .49$, with no difference between HBT and Waitlist at post. Similarly, there was a pre to post trend for decreasing physical activity in the HBT condition over time, $t(70.23) = -1.81$, $p = .075$, $d = -.38$.

A significant time by condition interaction was also found for IPAQ sedentary time, $F(4,66.08) = 2.56$, $p = .046$, which appeared due to participants in both the HBT+ACT and HBT condition reporting lower sedentary time than waitlist over time. IPAQ sedentary time was

significantly higher (i.e., worse) at post in the waitlist condition relative to both the HBT+ACT, $t(53.41) = 2.45$, $p = .018$, $d = .73$, and HBT conditions, $t(60.01) = 2.47$, $p = .017$, $d = .78$. The HBT+ACT and HBT conditions were equivalent on sitting time at post. There were significant pre to post decreases on sedentary time within both the HBT+ACT, $t(74.71) = 4.30$, $p < .001$, $d = .84$, and HBT conditions, $t(74.13) = 4.41$, $p < .001$, $d = .93$, but not the waitlist condition.

A final significant time by condition effect was found for TFEQ cognitive restraint, $F(4,70.47) = 2.52$, $p = .049$, which appeared due to cognitive restraint with eating failing to improve in the HBT condition. TFEQ cognitive restraint significantly improved from pre to post in both the HBT+ACT, $t(75.62) = 3.36$, $p = .001$, $d = .66$, and waitlist conditions, $t(75.66) = 2.49$, $p = .015$, $d = .48$, but not the HBT condition. There were no differences between conditions at post.

No time by condition effects were found for other outcome measures including WCCS weight control behaviors, TFEQ unrestrained and emotional eating, and GHQ mental health ($p > .10$). Similarly, no time by condition effects were found for the VLQ or CAQ valued action process measures ($p > .10$).

Based on previous findings that ACT Matrix app effects were stronger among participants who maintained active use of the app (Levin et al., 2017), post hoc analyses repeated MMRM tests with the subsample of participants who completed the evening check-in on at least half of the 28 day intervention period in the HBT ($n = 19$) and HBT+ACT conditions ($n = 24$). Results were nearly identical as with the full intent-to-treat sample, with time by condition effects only found for IPAQ physical activity, IPAQ sedentary time, and TFEQ cognitive restraint.

GLMM testing increases in valued action in HBT+ACT app data

GLMM was used to test the hypothesis that the rate of engaging in toward moves relative to away moves would increase with further exposure to the app (operationalized as the number of days since first downloading the app) among those in the HBT+ACT condition. The sample for this model was limited to those who were assigned to the HBT+ACT condition and completed at least one toward/away check-in (n = 29). Days since downloading the app significantly predicted the probability of engaging in a toward move (b = 0.05, odds ratio = 1.05, $p < 0.001$). This means that for each day, the probability of engaging in a toward move relative to an away move increased by 5 percent.

Discussion

This study sought to evaluate the acceptability and efficacy of self-monitoring avoidant and valued functions of behaviors using the ACT Matrix app relative to monitoring only the topography of health behaviors with a simple tracking app. This newest version of the ACT Matrix app received high usability ratings, but somewhat low ratings on other satisfaction variables, suggesting the app was easy to use, but had modest perceived helpfulness. Program usage was also mixed, with half of the prompted check-ins being completed on average, and 14% never using the ACT app. Participants in the ACT Matrix condition reported higher physical activity at post relative to the HBT and waitlist conditions, potentially due to the ACT Matrix app protecting against a decrease in physical activity over time observed in the other conditions. Yet, results were generally mixed on other health behavior outcomes and values processes with mostly null results relative to one or both of the other study conditions. Overall, findings suggest some benefits of the ACT Matrix app for physical activity, but with mixed findings on acceptability, outcomes, and processes of change suggesting results are preliminary and may be relatively limited.

Some of the results suggest the potential added benefits of self-monitoring the avoidant and valued functions of health behaviors. Most notably, participants assigned to the ACT Matrix app reported higher physical activity over time relative to the other two conditions. There was a non-significant trend for physical activity decreasing in the HBT condition over time, suggesting potentially that the ACT Matrix app protected against this decline in health behavior change success that can naturally occur following the initiation of a program as motivation dwindles and ineffective change efforts are employed. The ACT Matrix app also improved cognitive restraint with eating relative to the HBT app and sedentary time relative to waitlist.

The ACT Matrix has been hypothesized to increase psychological flexibility and valued living through several procedures (Polk & Schoendorff, 2014). The main procedure that was tested as a potential predictor of change in this study is noticing and labeling the function of one's observable behavior as either avoidance of unwanted internal experiences, or approach toward personal values. Valued behavior is expected to increase as individuals engage in this discrimination task more over time, due to noticing the function of their behavior (whether it is "about" avoidance or values), noticing when avoidant behavior is ineffective or inconsistent with values, and drawing more connections between personal values and behavior (making engaging in valued behavior more reinforcing; Polk & Schoendorff, 2014).

From this theoretical perspective, it is interesting that positive results were found for the ACT Matrix app in the areas of physical activity and eating-related cognitive restraint. It is possible that this is related to the function of these behaviors initially. For example, physical activity may be largely regulated by avoidance prior to intervention (e.g., exercising in an attempt to reduce body dissatisfaction), but relevant to personal values if clarified, such that it is particularly well-situated to change through this type of discrimination task. That is, participants

may have exercised at baseline as an away move, but increasingly noticed that exercising serves as a toward move, making exercise more rewarding and ultimately more frequent and consistent. This would be consistent with the trending decline in physical activity observed in the HBT condition. It may be that participants were motivated and engaged in increasing physical activity at baseline, but as a form of experiential avoidance, which led to a decrease in activity over time due to the maladaptive effects of this class of behaviors. Consistent with this, the majority of the sample reported high physical activity at baseline, with all but one of the participants engaging in at least moderate physical activity. In part, the ACT Matrix app may function to harness the efforts and motivation of individuals initiating physical activity change, while shifting the function from avoidance to values-based behaviors that are sustainable over time. If future studies replicate these results, it would be worthwhile to further examine potential moderators of treatment effects such as degree of avoidance and values clarity at baseline in different life domains. Future research would also benefit from testing the matrix app with sedentary samples that are not physically active.

These results suggest that incorporating this type of discrimination activity, or more broadly increasing awareness of avoidance and values may be particularly beneficial for interventions targeting physical activity, and possibly sedentary behavior and intentional regulation of eating. This awareness could be targeted through self-help resources (e.g., mindfulness apps) or by treatment providers teaching clients to regularly practice this type of discrimination.

Although there were some positive results, other results suggest that the ACT Matrix app, as well as the more basic health behavior tracking app, had limited effects. This included a lack of improvements on the WCCS health behavior measure, which was found to improve in the

initial pilot trial of the ACT Matrix app (Levin et al., 2017) and would be expected to improve from self-monitoring diet/exercise, as well as measures of emotional eating, uncontrolled eating, general psychological distress, and valued action that were expected to improve from the ACT Matrix app. The lack of improvements in several measures of diet/physical activity with the basic health behavior tracking app was surprising given previous research indicating the positive effects of self-monitoring (Korotitsch & Nelson-Gray, 1990; Lieffers & Hanning, 2012; Michie et al., 2009). This may indicate methodological issues such as decreased power due to the heterogeneous, non-clinical sample or weaknesses with the global self-report assessment methods used. These null results may also be due to the simplified self-monitoring approach used, which aimed to reduce response effort by tracking broad categories (e.g., healthy versus unhealthy meals), rather than the more time intensive, detailed tracking methods typically used. People are notably inaccurate in self-reporting their dietary intake (Archer, Pavela & Lavie, 2015) and physical activity (Lee et al., 2011; Sallis & Saelens, 2000), which may have been exacerbated with the simple tracking approach used that excluded more detailed caloric or activity tracking. These results suggest that self-monitoring diet and physical activity using this simplified, broad method may have limited efficacy.

Similarly, the revisions to the ACT Matrix app for the current study may have reduced its efficacy. For example, the initial pilot study included an in-person orientation, which was replaced with a self-guided online orientation to improve experimental control and generalizability (i.e., testing effects of an online intervention without in-person contact). The lack of an in-person orientation may have reduced efficacy, which is broadly consistent with existing literature indicating greater efficacy with online behavioral technologies when including personal contact (e.g., Spek et al., 2007). The ACT Matrix app included the same random prompts to self-

monitor the avoidant/valued functions throughout the day as the original pilot study (Levin et al., 2017). However, the revised app included additional evening and morning check-ins to further enhance awareness and monitoring of these functions in relation to health behaviors based on previous findings that participants desired and may obtain greater results from such additional features (Krafft et al., 2019; Levin et al., 2017). The current study results suggest these additions did not lead to substantially improved outcomes and may have even reduced effects given previous pilot findings. It is particularly noteworthy that the ACT Matrix app did not lead to improvements in valued action relative to the waitlist given this is the primary target and process of change that would be expected to improve from tracking behaviors in relation to personal values. That said, there was significant improvement in the ACT app condition on the rate of toward moves during daily life over time, suggesting that individuals may perceive overall values consistency differently from the moment-by-moment categorization of toward and away moves.

The ACT matrix app had mixed support for acceptability in the current study, which may have also contributed to the lack of effects on key outcomes and processes of change. The current version of the ACT matrix app appeared to have similar engagement rates to other mobile app studies (Krafft et al., 2019; Levin et al., 2017; Linardon & Fuller-Tyszkiewicz), but these rates were far from ideal with a portion of users never using the app at all and users adhering to only about half of the prompted check-ins. It is not clear what an adequate “dosage” is for the ACT matrix app to have its intended effects, or for other mobile apps more broadly, and thus it is not clear what benchmark to use to define acceptable engagement. Yet, these findings at least suggest users generally did not adhere to the full set of prompted check-ins they were expected to complete. Although satisfaction ratings were equivalent to prior evaluations of

the matrix app (Krafft et al., 2019; Levin et al., 2017), they were lower than other ACT mobile app ratings (Levin et al., 2019) and tended to fall below the face valid benchmark of 4 (*slightly agree*). Thus, it seems that revisions made to the current app (e.g., contextualizing toward and away more in relation to diet and physical activity, adding more exercises such as a morning and evening check-in, adding email summaries) did not improve its acceptability to participants. It may be that participants still desired a more complex, sophisticated app, particularly in terms of providing direct skills training rather than simply self-monitoring avoidant and valued functions. In other words, participants may desire, and benefit from, interventions that help them learn how to be less avoidant and take more valued actions, rather than interventions that only raise awareness of these patterns. This may have contributed to the modest engagement rates and the lack of effects on several outcome and process measures.

There are some notable limitations with the current study. The study had minimal inclusion criteria in order to mirror the broad range of users who might seek to use a mobile app to improve diet/exercise. This led to a heterogeneous sample without clearly identified problems with diet or physical activity. Participants likely varied widely in their goals for modifying diet and physical activity, but no validated measures were used to characterize differences and types of goals for using the app and changing health behaviors. Furthermore, the majority of the sample reported high physical activity at baseline. The increased variability and potential ceiling effects on diet and physical activity outcomes due to a lack of inclusion criteria might have reduced sensitivity to detecting treatment effects or generalizability to clinical samples. In addition, the sample was predominantly younger, White, and female, which may have further reduced generalizability to more diverse populations.

The study relied on global self-report measures of physical activity and dietary behaviors, which have limited validity relative to more objective, interview, or diary-based assessment methods. Although the study included a longer assessment window relative to the initial pilot trial (Levin et al., 2017), four weeks was still relatively short for assessing changes in diet and physical activity. This short time frame was used to focus on the immediate, initial effects of a mobile app, hypothetically when individuals are most actively motivated and engaged in using the app to make behavior changes. This also appeared particularly relevant given the simplicity of the self-monitoring apps being tested, which may not have been adequately sophisticated for more long term use. However, this short time frame may have reduced sensitivity to detecting the full effects of the self-monitoring app. Furthermore, the study did not include an assessment of long-term follow up effects to determine if improvements were maintained over time, which is particularly relevant to health behaviors such as diet and physical activity.

The two active conditions were not balanced on the amount of check-ins and app exercises, which introduced an alternate explanation for any differences found between conditions. The comparison app might have been just as effective as HBT+ACT on physical activity and cognitive restraint if it had included a morning check-in and random prompts throughout the day. There were also failures of randomization on variables including committed action and BMI between conditions. Although these were addressed by including these variables as covariates, this might have been addressed proactively through stratified randomization on key variables.

Overall, this study adds to a growing literature on the potential benefits of ACT for improving diet and physical activity and ACT apps more specifically. The results were mixed, but provided some indication that tracking avoidant and valued functions of behavior through the

ACT Matrix app can improve health behaviors, particularly physical activity. Future self-monitoring apps may benefit from including tools to track the function of behavior in addition to topography.

References

Archer, E., Pavela, G. & Lavie, C.J. (2015). The inadmissibility of What We Eat in America and NHANES dietary data in nutrition and obesity research and the scientific formulation of national dietary guidelines. *Mayo Clin Proc, 90,* 911-926.

Bandura, A. (1998). Health promotion from the perspective of social cognitive theory. *Psychology and Health, 13,* 623–649. http://doi.org/10.1080/08870449808407422

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction, 24,* 574–594.

Banks, M. H. (1980). The use of the general health questionnaire as an indicator of mental health in occupational studies. *Journal of Occupational Psychology, 53,* 187-194.

Booth, M. L. (2000). Assessment of physical activity: An international perspective. *Research Quarterly for Exercise and Sport, 71,* s114-120.

Bricker, J.B., Mull, K., Kientz, J.A., Vilardaga, R.M., Mercer, L.D., Akioka, K. & Heffner, J.L. (2014). Randomized controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. *Drug and Alcohol Dependence, 143,* 87-94.

Brooke, J. (1996). SUS: a "quick and dirty" usability scale. In P.W. Jordan, B. Thomas, B.A. Weerdmeester, and I.L. McClelland (Eds.) *Usability Evaluation in Industry* (189-194). London: Taylor and Francis.

Carter, M. C., Burley, V. J., Nykjaer, C., & Cade, J. E. (2013). Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *Journal of Medical Internet Research, 15,* e32. http://doi.org/10.2196/jmir.2283

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., ... & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, *35*, 1381-1395.

de Lauzon, B., Romon, M., Deschamps, V., Lafay, L., Borys, J.-M., Karlsson, J., … Fleurbaix Laventie Ville Sante (FLVS) Study Group. (2004). The Three-Factor Eating Questionnaire-R18 is able to distinguish among different eating patterns in a general population. *The Journal of Nutrition*, *134*, 2372–2380. https://doi.org/10.1093/jn/134.9.2372

Forman, E. M., & Butryn, M. L. (2015). A new look at the science of weight control: how acceptance and commitment strategies can address the challenge of self-regulation. *Appetite*, *84*, 171-180.

Forman, E. M., Butryn, M. L., Manasse, S. M., Crosby, R. D., Goldstein, S. P., Wyckoff, E. P. & Thomas, J. G. (2016). Acceptance-based versus standard behavioral treatment for obesity: Results from the mind your health randomized controlled trial. *Obesity, 24*, 2050-2056.

Gatewood, J. G., Litchfield, R. E., Ryan, S. J., Myers Geadelmann, J. D., Pendergast, J. F., & Ullom, K. K. (2008). Perceived barriers to community-based health promotion program participation. *American Journal of Health Behavior*, *32*, 260–271.

Goldberg, D. P. (1972). The detection of psychiatric illness by questionnaire. London, UK: Oxford University Press.

Hagströmer, M., Oja, P., & Sjöström, M. (2006). The International Physical Activity Questionnaire (IPAQ): A study of concurrent and construct validity. *Public Health Nutrition, 9,* 755-762.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and Commitment Therapy: The process and practice of mindful change*. New York: The Guilford Press.

Karlsson, J., Persson, L.-O., Sjöström, L., & Sullivan, M. (2000). Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study. *International Journal of Obesity, 24,* 1715–1725. https://doi.org/10.1038/sj.ijo.0801442

Korotitsch, W. J. & Nelson-Gray, R. O. (1990). An overview of self-monitoring research in assessment and treatment. *Psychological Assessment, 11,* 415-425.

Krafft, J., Potts, S., Schoendorff, B., & Levin, M. E. (2017). A randomized controlled trial of multiple versions of an acceptance and commitment therapy matrix app for well-being. *Behavior Modification, 43,* 246–272. https://doi.org/10.1177/0145445517748561

Lee, P. H., Macfarlane, D. J., Lam, T. H., & Stewart, S. M. (2011). Validity of the international physical activity questionnaire short form (IPAQ-SF): A systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 8,* 115-126.

Levin, M.E., Haeger, J. & Cruz, R.A. (2019). Tailoring acceptance and commitment therapy skill coaching in-the-moment through smartphones: Results from a randomized controlled trial. *Mindfulness, 10,* 689-699.

Levin, M.E., Pierce, B. & Schoendorff, B. (2017). The acceptance and commitment therapy matrix mobile app: A pilot randomized trial on health behaviors. *Journal of Contextual Behavioral Science, 6,* 268-275.

Lieffers, J. R. L., & Hanning, R. M. (2012). Dietary assessment and self-monitoring: With nutrition applications for mobile devices. *Canadian Journal of Dietetic Practice and Research, 73,* 253–260. http://doi.org/10.3148/73.3.2012.e253

Lillis, J., Hayes, S. C. & Levin, M. E. (2011). Binge eating and weight control: The role of experiential avoidance. *Behavior Modification, 35*, 252-264.

Lillis, J. & Kendra, K. E. (2014). Acceptance and commitment therapy for weight control: Model, evidence, and future directions. *Journal of Contextual Behavioral Science, 3*, 1-7.

Linardon, J. (in press). Can acceptance, mindfulness, and self-compassion be learnt by smartphone apps? A systematic and meta-analytic review of randomized controlled trials. *Behavior Therapy*.

Linardon, J., & Fuller-Tyszkiewicz, M. (In Press). Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of Consulting and Clinical Psychology.*

Litwin, R., Goldbacher, E. M., Cardaciotto, L. & Gambrel, L. E. (2017). Negative emotions and emotional eating: The mediating role of experiential avoidance. *Eating and Weight Disorders, 22*, 97-104.

Loef, M., & Walach, H. (2012). The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. *Prev Med, 55*, 163-170.

McCracken, L. M., Chilcot, J. & Norton, S. (2014). Further development in the assessment of psychological flexibility: A shortened committed action questionnaire (CAQ-8). *European Journal of Pain, 19*, 677-685.

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychology, 28*, 690–701. http://doi.org/10.1037/a0016136

Moffitt, R., & Mohr, P. (2015). The efficacy of a self-managed acceptance and commitment therapy intervention DVD for physical activity initiation. *British Journal of Health Psychology, 20,* 115-129.

Mozaffarian, D., Wilson, P. W., & Kannel, W. B. (2008). Beyond established and novel risk factors: lifestyle risk factors for cardiovascular disease. *Circulation, 117*, 3031-3038.

Niemeier, H. M., Leahey, T., Reed, K. P., Brown, R. A. & Wing, R. R. (2012). An acceptance-based behavioral intervention for weight loss: A pilot study. *Behavior Therapy, 43*, 427-435

Palmeira, L., Cunha, M., & Pinto-Gouveia, J. (2018). The weight of weight self-stigma in unhealthy eating behaviours: The mediator role of weight-related experiential avoidance. *Eating and Weight Disorders, 23*, 785-796.

Payne, H. E., Lister, C., West, J. H., & Bernhardt, J. M. (2015). Behavioral functionality of mobile apps in health interventions: A systematic review of the literature. *JMIR MHealth and UHealth, 3*, e20. http://doi.org/10.2196/mhealth.3335

Peng, W., Kanthawala, S., Yuan, S., & Hussain, S. A. (2016). A qualitative study of user perceptions of mobile health apps. *BMC Public Health, 16,* 1–11. http://doi.org/10.1186/s12889-016-3808-0

Polk, K. L., & Schoendorff, B. (2014). *The act matrix: A new approach to building psychological flexibility across settings and populations*. Context Press.

Pinto, A. M., Fava, J. L., Raynor, H. A., LaRose, J. G., & Wing, R. R. (2013). Development and validation of the weight control strategies scale. *Obesity, 21,* 2429-2436.

Reiner, M., Niermann, C., Jekauc, D., & Woll, A. (2013). Long-term health benefits of physical activity: A systematic review of longitudinal studies. *BMC Public Health, 13*, 1–9. http://doi.org/10.1186/1471-2458-13-813

Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: status, limitations, and future directions. *Research Quarterly for Exercise and Sport*, *71*, 1-14.

Salmon, J., Crawford, D., Owen, N., Bauman, A., & Sallis, J. F. (2003). Physical activity and sedentary behavior: A population-based study of barriers, enjoyment, and preference. *Health Psychology*, *22*(2), 178–188. http://doi.org/10.1037/0278-6133.22.2.178

Spek, V., Cuijpers, P. I. M., Nyklíček, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: A meta-analysis. *Psychological Medicine, 37*, 319-328.

Teixeira, P. J., Carraca, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 9*, 78-108.

Torous, J. B., Levin, M. E., Ahern, D., & Oser, M. (2017). Cognitive behavioral mobile applications: Research literature, marketplace data, and evaluation guidelines. *Cognitive and Behavioral Practice, 24*, 215-225.

Ulmer, C. S., Stetson, B. A. & Salmon, P. G. (2010). Mindfulness and acceptance are associated with exercise maintenance in YMCA exercisers. *Behaviour Research and Therapy, 48*, 805-809.

Vestuyf, J., Patrick, H., Vansteenkitse, M., & Teixeira, P.J. (2012). Motivational dynamics of eating regulation: a self-determination theory perspective. *International Journal of Behavioral Nutrition and Physical Activity*, 9.

Wang, D. D., Li, Y., Chiuve, S. E., Hu, F. B., & Willett, W. C. (2015). Improvements In US Diet Helped Reduce Disease Burden And Lower Premature Deaths, 1999-2012; Overall Diet Remains Poor. Health Aff (Millwood), 34(11), 1916-1922.

Wilson, K. G., Sandoz, E. K., Kitchens, J. & Roberts, M. (2010). The valued living questionnaire: Defining and measuring valued action within a behavioral framework. *The Psychological Record, 60*, 249-272

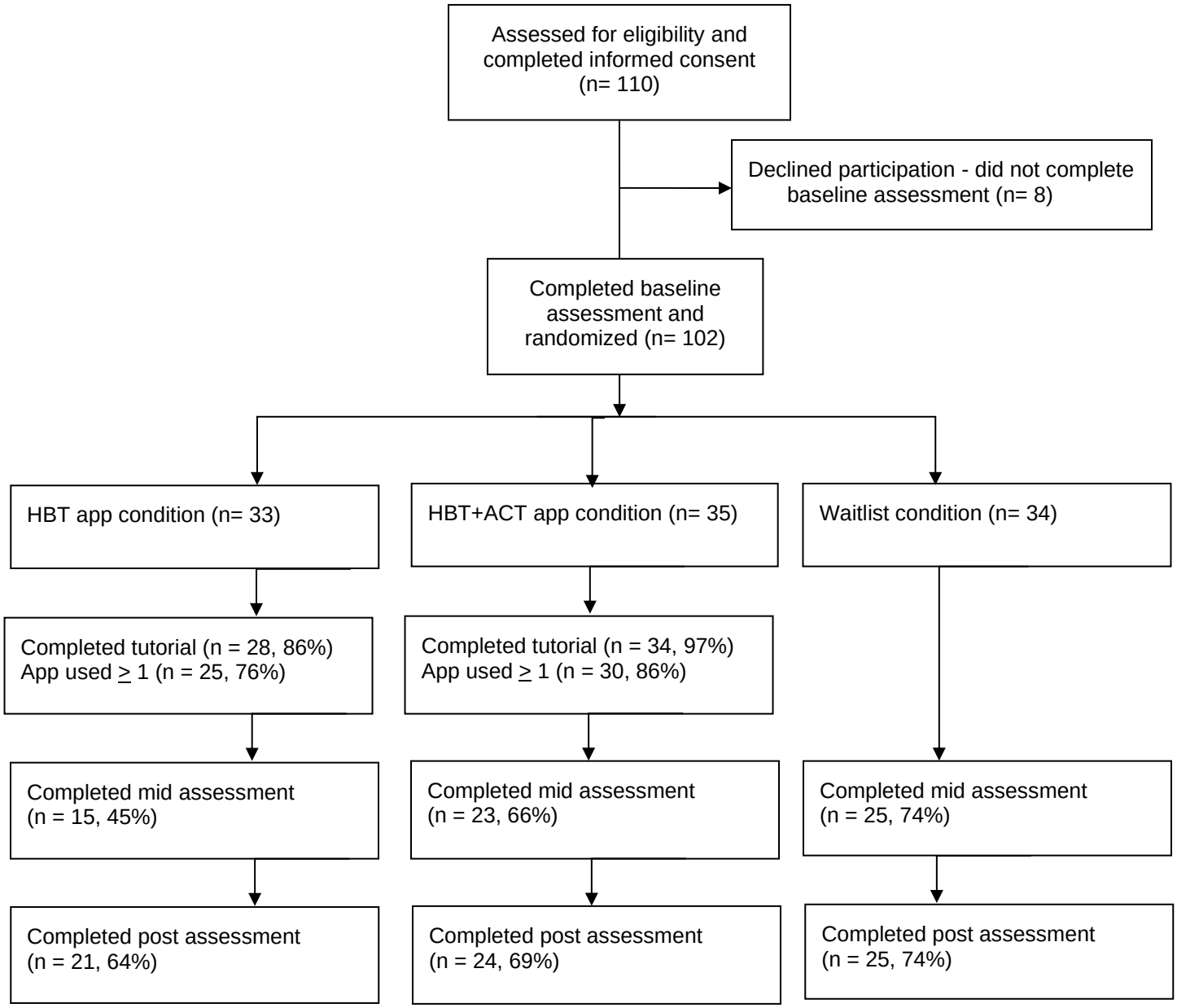
Table 1. Estimated descriptive statistics by condition and time point based on MMRM analyses.

	HBT+ACT App			HBT App			Waitlist		
	Pre *M*(*SE*)	Mid *M*(*SE*)	Post *M*(*SE*)	Pre *M*(*SE*)	Mid *M*(*SE*)	Post *M*(*SE*)	Pre *M*(*SE*)	Mid *M*(*SE*)	Post *M*(*SE*)
Outcome Measures									
IPAQ-Total	6919.71 (1587.67)	6313.12 (1321.64)	8314.64 (1550.40)	6796.40 (1581.00)	6311.40 (1374.28)	5122.90 (1577.38)	6279.08 (1494.04)	3384.69 (1172.27)	5508.97 (1449.02)
IPAQ-Sed	2778.71 (336.43)	2135.98 (279.78)	1793.79 (291.02)	2894.36 (337.18)	1973.94 (307.03)	1928.65 (302.45)	2836.99 (319.03)	2538.49 (248.45)	2651.30 (264.74)
WCCS	68.63 (3.61)	82.97 (3.83)	80.55 (4.10)	69.78 (3.56)	80.44 (4.19)	83.39 (4.22)	67.13 (3.30)	73.55 (3.43)	74.04 (3.81)
TFEQ-CR	14.08 (.76)	15.02 (.69)	15.77 (.75)	14.00 (.75)	14.02 (.72)	14.35 (.76)	13.83 (.70)	15.61 (.75)	15.06 (.68)
TFEQ-UE	22.20 (1.28)	21.04 (1.33)	20.38 (1.37)	22.26 (1.26)	21.51 (1.39)	20.94 (1.37)	19.99 (1.16)	20.10 (1.19)	19.13 (1.25)
TFEQ-EE	7.81 (.54)	7.21 (.58)	7.14 (.56)	7.70 (.53)	6.87 (.60)	7.23 (.56)	7.29 (.49)	6.93 (.52)	7.19 (.50)
GHQ	27.39 (1.15)	22.75 (1.24)	22.74 (1.14)	27.50 (1.14)	23.93 (1.44)	24.74 (1.18)	26.33 (1.07)	25.04 (1.11)	24.70 (1.05)
Process Measures									
VLQ	5281.07 (367.80)	5175.20 (414.33)	5706.09 (421.00)	5129.45 (360.19)	5406.11 (435.20)	5138.03 (433.65)	5603.29 (331.13)	5480.46 (377.77)	5747.06 (38.83)
CAQ	34.08 (1.21)	35.70 (1.23)	36.47 (1.34)	38.70 (1.26)	39.04 (1.41)	39.61 (1.44)	35.25 (1.22)	35.98 (1.21)	35.28 (1.32)

Descriptive statistics estimated in MMRM with the full ITT sample and controlling for baseline differences on the CAQ and BMI. CAQ analysis only included BMI as a covariate. Raw scores are provided in this table for variables that were transformed for MMRM analyses (IPAQ-Total, IPAQ-Sed, GHQ). IPAQ-Total = International Physical Activity Questionnaire Total score; IPAQ-Sed = IPAQ Sedentary Time; WCCS = Weight Control Strategies Scale; TFEQ-CR = Three-Factor Eating Questionnaire-Cognitive Restraint; TFEQ-UE = TFEQ Uncontrolled Eating; TFEQ-EE = TFEQ Emotional Eating; GHQ = General Health Questionnaire; VLQ = Valued Living Questionnaire; CAQ = Committed Action Questionnaire.

Figure caption

Figure 1. Participant flow diagram.


Assessed for eligibility and completed informed consent (n= 110)
Declined participation - did not complete baseline assessment (n= 8)
Completed baseline assessment and randomized (n= 102)
HBT app condition (n= 33)
HBT+ACT app condition (n= 35)
Waitlist condition (n= 34)
Completed tutorial (n = 28, 86%)
App used ≥ 1 (n = 25, 76%)
Completed tutorial (n = 34, 97%)
App used ≥ 1 (n = 30, 86%)
Completed mid assessment (n = 15, 45%)
Completed mid assessment (n = 23, 66%)
Completed mid assessment (n = 25, 74%)
Completed post assessment (n = 21, 64%)
Completed post assessment (n = 24, 69%)
Completed post assessment (n = 25, 74%)

‘