Received: 6 October 2021 | Revised: 24 February 2022 | Accepted: 24 February 2022

DOI: 10.1002/eat.23700

ORIGINAL ARTICLE

International Journal of EATING DISORDERS WILEY

Understanding self-monitoring to inform a mobile intervention for binge eating and weight management: A proof-of-concept randomized trial

Jianyi Liu MA[1,2] | Sean A. Munson PhD[3] | Angela Chang BA[1,4] | Claire Voss BA[1] | Andrea K. Graham PhD[1,4]

[1]Center for Behavioral Intervention Technologies, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

[2]Department of Psychiatry & Behavioral Science, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

[3]Human Centered Design & Engineering, University of Washington, Seattle, Washington, USA

[4]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

Correspondence
Andrea Graham, PhD, 750 N Lake Shore Dr, 10th Floor, Chicago, IL 60642, USA.
Email: andrea.graham@northwestern.edu

Funding information
National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: K01 DK116925

Action Editor: Tracey Wade

Abstract

Objective: This study explored consumers' perspectives on self-monitoring, a common feature in behavioral interventions that helps inform consumers' progress and answer their questions, to learn what outcome metrics matter to consumers and whether self-selection of these metrics leads to greater engagement (i.e., compliance, satisfaction) in self-monitoring than monitoring only default options.

Methods: In a proof-of-concept randomized trial, 48 adult participants were randomly assigned to "clinician-determined monitoring" or "clinician + self-determined monitoring" conditions. Before starting monitoring, all participants shared outcomes that would matter to them in a mobile intervention for binge eating and weight management. Then, for 3 weeks, participants in the "clinician-determined" condition monitored their weight and binge-eating episodes, and participants in the "clinician + self-determined" condition monitored these and another metric of their choosing. After, satisfaction and compliance were assessed.

Results: Participants identified 116 metrics, grouped into 12 themes, that mattered to them. During monitoring, participants in the "clinician + self-determined" condition monitored 41 metrics. Surprisingly, participants in the "clinician-determined" condition also monitored metrics besides weight and binge eating. This resulted in a failure of our experimental manipulation, which represents a significant limitation of this research. No significant differences emerged in satisfaction or compliance between conditions.

Discussion: Although our proof-of-concept trial yielded null quantitative results, findings also suggested binge eating and weight management interventions may benefit from including an individually customizable monitoring option in addition to default metrics, warranting testing in future research.

Public Significance: Examining consumers' self-monitoring preferences for a mobile intervention for binge eating and weight management revealed a variety of metrics

This study was previously presented at the 2021 Academy for Eating Disorders International Conference on Eating Disorders, June 10, 2021.

that matter to consumers, although binge eating and weight were still most valued. Findings from our proof-of-concept trial suggest design implications of encouraging an individually customizable monitoring option, in addition to default metrics, which needs to be tested in future research over a longer period and during actual mobile intervention delivery.

KEYWORDS
binge eating, mobile intervention, self-monitoring, user-centered design, weight management

1 | INTRODUCTION

Self-monitoring, in which a person regularly observes and records their behavior(s) or the outcomes of behavior(s) within a behavior change strategy, is an evidence-based technique for behavior change and a common feature of many evidence-based interventions for mental and behavioral health problems (Michie et al., 2013). While this technique can be burdensome to people engaged in interventions ("consumers"), it can inform both clinicians and consumers about consumers' progresses (Fortney et al., 2017; Shimokawa et al., 2010; Torous & Roux, 2017; J. P. Tregarthen et al., 2015). In this paper, "consumer" refers to those who engage in interventions (e.g., using a mobile app). Other terms such as "users" have been associated with infantilizing and oversimplifying people, and "patients" may not be suitable for all experiences.

Consumers and clinicians can have different interests regarding what metrics consumers should monitor in an intervention, as what matters to them can differ. Qualitative analyses of open-ended responses from patients with a history of depression, informal caregivers, and health care professionals revealed 137 different outcomes domains that are of value to these individuals, 80 of which related to the benefits of treatment (Chevance et al., 2020). For example, the authors identified 19 domains related to mood and emotional symptoms (e.g., relief from mental pain, anxiety, sadness) and 18 domains associated with cognitive symptoms (e.g., motivation, cognitive distortion, and social interest). These distinctive outcomes extend beyond standard depression assessments, and the authors noted that clinical trials rarely assess for many of these domains (Chevance et al., 2020). Additionally, in another study that investigated patients' perspectives on benefits from a depression intervention, 20 participants reported 6 themes and 15 subthemes of personal benefits that matter to themselves (Folkersma et al., 2021). These studies suggest there might be a wide range of meaningful metrics from consumers' perspectives, making it important to examine what metrics matter to consumers. Research also has demonstrated that a large proportion of people who engage with behavior change interventions want customizable options for monitoring, different from the default options requested by clinicians (Karkar et al., 2017; Schroeder et al., 2018).

Understanding these differences is important because not attending to the differences can negatively impact consumers' engagement with an intervention (Munson et al., 2020). In one study of digital wellness tracking programs in the workplace, not having customizable tracking options was a main reason why some participants decided not to use the app or stopped using the app (Chung et al., 2017). These participants indicated that if they could not track metrics that fit with their health goals, this program might not be the best choice. A separate study found that different interests in tracking goals and tools used in medical care settings could result in tension between patients and clinicians and contribute to patient disengagement (Chung et al., 2016).

Conversely, having default monitoring options that are recommended by clinicians can be useful. While some patients have preferences for what they will track, others look to their health clinicians for guidance on what data are relevant and likely to lead to actionable insights (Chung et al., 2016; Lordon et al., 2020). In a study evaluating designs of a self-tracking app for patients with migraines, participants wanted the app to recommend default options for monitoring, in addition to being able to add their own. Participants indicated that they might not choose the default option, but they appreciated the guidance provided when deciding what to monitor (Schroeder et al., 2019).

Research in eating disorders has not yet focused on understanding what outcomes matter to consumers to monitor in an intervention. Such an examination is necessary because self-monitoring is a core component of some evidence-based interventions like cognitive-behavioral therapy (CBT) for eating disorders (Fairburn, 2008), and completing self-monitoring can have clinical benefits (Latner & Wilson, 2002). Yet, barriers can prevent monitoring or thwart sustaining monitoring (e.g., Cordeiro et al., 2015). Consequently, it is useful to understand how to make this technique most impactful and engaging to consumers. Additionally, previous research on a smartphone intervention with self-monitoring for eating disorders (Kim et al., 2018) found that consumers who received a personalized version with content tailored to their baseline symptoms showed significantly better improvement in their Eating Disorder Examination Questionnaire scores compared to consumers who received the standardized CBT-based version (J. Tregarthen et al., 2019). Therefore, having personalized content in a digital intervention with self-monitoring could benefit consumers, and it is important to investigate what outcome matter to consumers to design personalized interventions.

User-centered design (also referred to as human-centered design) is an approach that includes a myriad of methods for deeply understanding stakeholder needs and preferences and a given context to design compelling, effective, and engaging services and products. In

user-centered design, the end-user (i.e., target consumer of a product) is placed at the center of the design process to ensure that the resulting product meets the needs and preferences of those who engage with and/or are impacted by it (Graham et al., 2019; Lyon & Koerner, 2016; Norman, 1988; Norman & Draper, 1986). We applied user-centered design methods through a needs assessment and prototyping activity to inform the design of the self-monitoring component of a mobile intervention for binge eating and weight management (Graham, Munson, et al., 2021; Graham, Neubert, et al., 2021; Weinheimer et al., 2020). With the needs assessment, our aim was to learn what metrics consumers want to monitor in an intervention besides standard (i.e., default) options (study aim 1). With the prototyping activity, our aim (study aim 2) was to conduct a proof-of-concept randomized trial that tested the concept of whether monitoring customized metrics determined by the consumer versus only clinical defaults leads to differences in engagement (i.e., compliance, satisfaction)—the intended outcome of user-centered design. Our proof-of-concept trial hypothesis was that monitoring self-determined metrics would lead to higher user satisfaction and compliance, compared to only monitoring clinician-determined default metrics. To our knowledge, this is the first study to (1) identify metrics that are desirable to monitor among individuals seeking treatment for binge eating and weight management, and (2) test the impact of different monitoring strategies on design outcomes.

2 | METHODS

2.1 | Design

This study was a proof-of-concept randomized trial in which participants were randomly assigned to one of two study conditions. In the "clinician-determined monitoring" condition, participants were asked to monitor their weight and the number of binge-eating episodes they experienced for 3 weeks. In the "clinician + self-determined monitoring" condition, participants monitored their weight, the number of binge-eating episodes they experienced, and another metric of their choosing for 3 weeks. This study was approved by the Northwestern University Institutional Review Board, and the trial was registered at ClinicalTrials.gov (NCT04711577). All participants provided online informed consent.

2.2 | Participants

Participants ($N = 50$) were recruited on dscout, an online qualitative research platform. Participants were eligible if they were non-pregnant, English-speaking adults (age ≥ 18 years old) with self-reported obesity (body mass index ≥30 kg/m^2) and self-reported recurrent binge eating (≥12 binge-eating episodes over the past 3 months). Participants also needed to be interested in using an app to manage their weight and binge eating and to engage in self-monitoring; although this study did not deploy an intervention besides self-monitoring (see Section 2.3), these entry criteria aimed to ensure the sample reflected future consumers of a digital intervention and so findings could generalize to future implementation. Participants also were required to own or have access to a scale on a regular basis to measure their weight.

Among those eligible, the final sample was selected based on sex, age, race, and ethnicity to ensure a diverse group. Two participants (one in each condition) dropped out before submitting any monitoring entries, and therefore were excluded from the final analyses since they did not have any monitoring entries to assess, resulting in a final sample of 48 participants who initiated and completed the proof-of-concept trial.

2.3 | Procedure

2.3.1 | Recruitment and enrollment

All study procedures occurred via dscout: an online qualitative research platform with >100,000 users who can voluntarily respond to advertisements for qualitative research studies. The key feature of dscout is facilitating digital diary studies that enable users to report on experiences in the moments and contexts in which they occur, in response to prompts created by researchers. Dscout allows researchers to ask various types of questions, including multiple-choice questions, open-ended text questions, and media-based entries (i.e., videos, photographs).

This study used the dscout research platform to mimic self-monitoring; we did not create a monitoring app. No clinical intervention was delivered beyond asking participants to self-monitor (e.g., no clinical content was provided, and participants did not receive feedback on their self-monitoring entries). Questions asking participants to report on what they would like in a mobile intervention were intended to gauge their perspectives for a future digital intervention. Data S1 presents screenshots of a self-monitoring entry.

Participants were recruited via an advertisement on dscout for a paid opportunity to understand self-monitoring behaviors in mobile interventions for weight management and binge eating in a 1-month research study. Interested individuals provided online informed consent and then completed an online screening questionnaire to assess eligibility. Of the 821 individuals who initiated the screen, 50 eligible individuals were selected for and invited to participate. All invited individuals accepted the invitation to enroll. Figure 1 presents a consort diagram.

2.3.2 | Sample size and randomization

To facilitate two study arms, two parallel projects were conducted in dscout, one project per study condition. Dscout's sample size limit is 25 participants per project. Since this study was a proof-of-concept study (i.e., aimed to answer whether this research question is worth asking, not to estimate effect sizes), we decided that the dscout project cap would be sufficient.

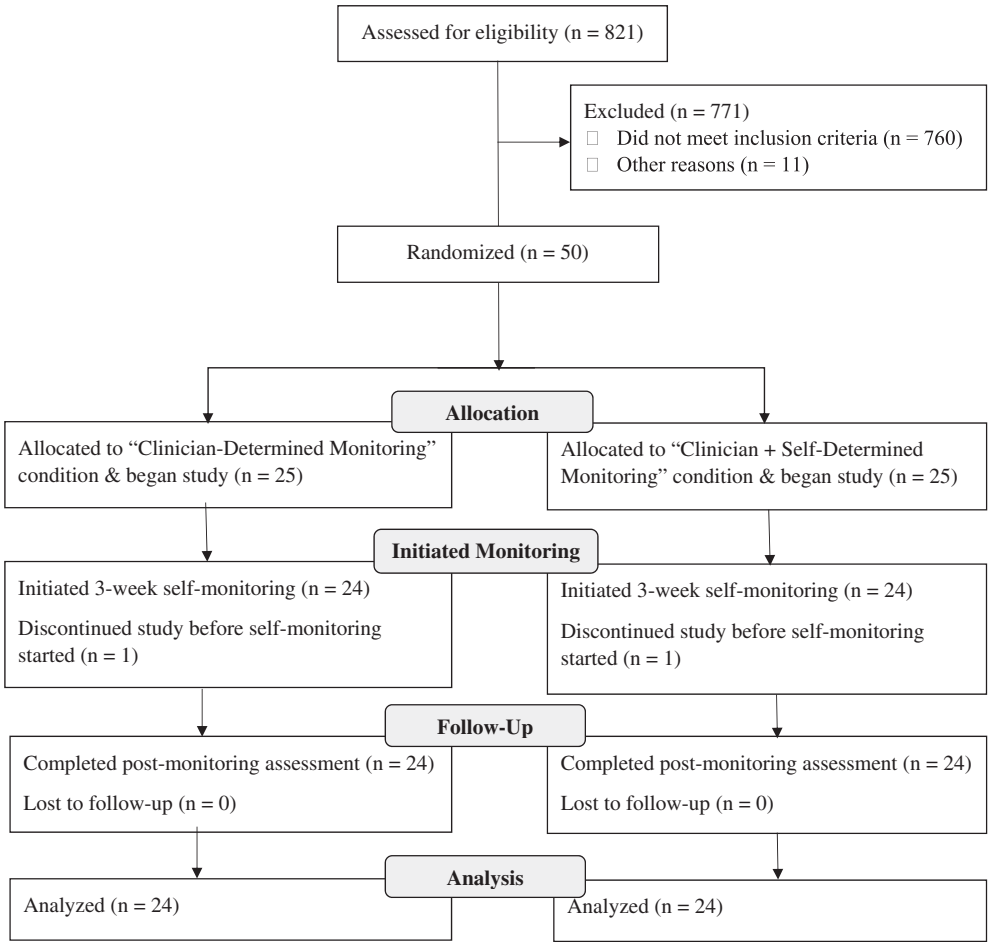

FIGURE 1 CONSORT participant flow diagram

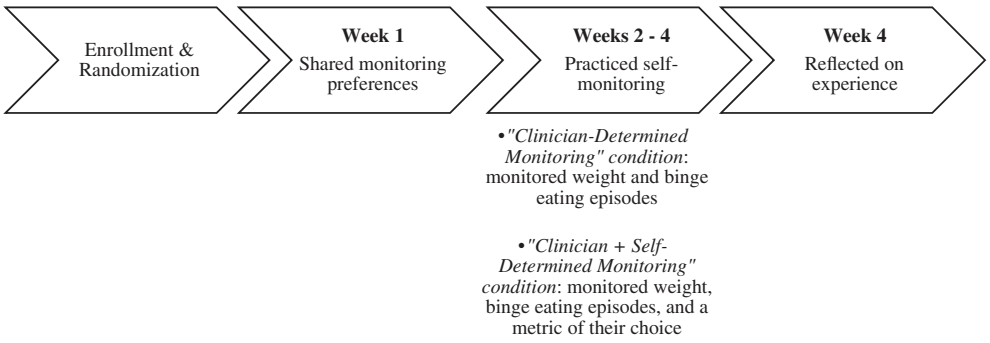

FIGURE 2 Study procedures

Participants were randomized into one of two study conditions—"clinician-determined monitoring" or "clinician + self-determined monitoring"—prior to the start of the project. Participants were randomized using a computer-generated sequence with a one-to-one ratio in blocks of four and six.

2.3.3 | Study procedure

Study procedures are presented in Figure 2. During the first week, all participants, across both study conditions, were asked to share a list of outcomes that would matter to them regarding self-monitoring, in addition to binge eating and weight, within a mobile intervention. They provided these data in a video recording and via an open-ended text response.

Participants were then invited to practice self-monitoring for 3 weeks. Each week, participants were asked to submit two self-monitoring entries (although they could submit more than one response for each entry) for a total of six requested entries. Entries were made available on the same days each week, such that participants had 3–4 days to complete each entry depending on when the entry was made available. Participants could submit their entry at any time during the requested window. Entries were made available sequentially so that participants would not submit all six entries at once. Participants received reminders before that entry's scheduled deadline. If the deadline passed and an entry still was not submitted, participants were asked to leave

that entry blank and move on to the next one. However, expired entries could not be locked, so it was technically possible for participants to still submit entries past the deadline. Previous research has shown that the ability to enter missing information can be important for sustaining tracking motivation (Cordeiro et al., 2015).

Participants' study condition differentiated what they were asked to self-monitor for the 3 weeks. For the first entry, participants in the "clinician-determined monitoring" condition monitored their weight and the number of binge episodes they had in the past week. Participants in the "clinician + self-determined monitoring" condition monitored those same metrics plus a metric of their choosing. These participants were encouraged to select something that they could monitor for the three-week duration. For each subsequent entry, monitoring instructions were the same except participants were asked to indicate the number of binge episodes that occurred since the previous entry. For each entry, participants were provided "fill-in-the-blank" (open-ended) response options for the metrics we requested of that study condition. Specifically, in the "clinician-determined monitoring" condition, participants were given two blank spaces to report their weight and number of binge episodes; in the "clinician + self-determined monitoring" condition, participants were given three blank spaces to report their weight, number of binge episodes, and the self-determined metric. After reporting these metrics, participants from both conditions completed a video recording to capture their qualitative experiences with monitoring, and then responded to an open-ended question asking if there was "anything else that would be useful for us to know."

After monitoring for 3 weeks, all participants were asked to reflect on their experience. To assess satisfaction with self-monitoring, participants completed the seven-item Satisfaction subscale of the Usefulness, Satisfaction, and Ease of Use (USE) Questionnaire (Lund, 2001), a validated measure that assesses usability (Gao et al., 2018). Items are rated on a scale of 1 (strongly disagree) to 7 (strongly agree), with higher scores indicating greater satisfaction. Also, participants indicated their satisfaction toward the frequency of monitoring (i.e., twice a week) by indicating if this frequency was "not enough," "just right," or "too much." Additionally, participants were asked to indicate which of the following metric(s) they found most important to monitor: weight, number of binge-eating episodes, nothing, something else (fill in), and the outcome they monitored over the past few weeks (assessed only among the "clinician + self-determined monitoring" condition). Participants could select multiple outcomes. Finally, participants were asked to indicate their likelihood of continuing to monitor without monetary reward on a scale of 1 (not at all) to 10 (absolutely).

Participants who completed all study procedures received a $75 reward.

2.4 | Outcomes

The main outcomes were differences between study conditions in compliance and self-reported satisfaction. Compliance was based on the number of monitoring entries completed of the expected six entries. Therefore, we defined compliance as the proportion of the sample who submitted at least six entries on time. Because individuals could submit multiple responses to each of the six entries, we also calculated the total number of entries submitted on time. Satisfaction was defined as the total score on the USE Questionnaire Satisfaction subscale.

Secondary outcomes included participant satisfaction toward monitoring frequency, perception of the most important metric(s) to monitor, and likelihood of continuing to monitor without monetary rewards.

2.5 | Analyses

Video recordings were automatically transcribed by dscout and were subsequently edited for accuracy and deidentification by an approved study team member. Transcripts from video recordings and open-ended text responses were combined for qualitative analyses, which were conducted to assess what metrics participants want to monitor, evaluated among the full sample before the monitoring period and evaluated separately by condition for responses during the monitoring period. Using methods from thematic analysis (Braun & Clarke, 2006, 2021), we used an approach that involved becoming familiar with participants' entry data and video transcriptions, coding data, and conceptualizing and summarizing codes into meaningful domains. Regarding reflexivity, the analysis was completed by an independent rater (psychology Master's student with research experience in eating disorders), who routinely reviewed and discussed the codes with another rater (clinical psychologist researcher with expertise in eating disorders and weight management) to reach consensus. The research team's collective expertise in clinical science, digital intervention, and human-computer interaction was used in understanding the results and developing the themes. The number of participants who mentioned and selected each outcome were counted.

The median of the satisfaction ratings for each condition were calculated, and difference in median satisfaction between conditions was evaluated using a Mann Whitney test. A Fisher's exact test was performed to compare differences in the proportion of the sample who was compliant. Descriptives are presented for the secondary outcomes. These analyses were conducted using SPSS version 26. p Values of <.05 were considered statistically significant.

3 | RESULTS

3.1 | Sample characteristics

Detailed demographic information for the full sample and for each condition is presented in Table 1. Participants were adults with a mean age of 38.12 years (SD = 10.86). Over half of participants

identified as White ($n = 27$; 58%), female ($n = 25$; 52%), and as college and post-college graduates ($n = 36$; 75%).

3.2 | What do people want to monitor?

Participants reported 116 distinct metrics that they were interested in monitoring (Data S2). As shown in Table 2, we grouped these metrics into 12 themes. The most common theme was metrics related to food and drink, comprised of 53 metrics indicated by 28 participants. The next most common theme was metrics related to physiology and body size (38 metrics indicated by 21 participants), followed by metrics related to physical activity (33 metrics indicated by 21 participants).

3.3 | What do people choose to monitor?

Table 3 shows the comprehensive list of what participants chose to monitor, by study condition, during the 3-week monitoring period. In the "clinician + self-determined monitoring" condition, participants monitored 41 different metrics. Two-fifths of participants in this condition ($n = 10$; 42%) chose to begin monitoring metrics related to food and drink. Over the 3 weeks, 50% of participants in this condition ($n = 12$) changed the metrics they were monitoring, despite instructions to monitor the same metric for 3 weeks.

In the "clinician-determined monitoring" condition, in which participants were instructed to only monitor weight and binge eating, nearly all participants ($n = 23$; 96%) indicated via their video recordings and/or open-ended text responses that they monitored additional metrics on their own besides their weight and binge episodes over the monitoring period. Of these individuals, 10 participants (42%) chose to monitor their food intake, six (25%) chose to monitor calorie intake, and the rest monitored different outcomes.

TABLE 2 Different themes that participants wanted to monitor

Theme	No. of metrics per theme	*n* (%[a]) indicated this theme
Food and drink	53	28 (58%)
Physiology and body size	38	21 (44%)
Physical activity	33	21 (44%)
Mental health	26	13 (27%)
Sleep	14	9 (19%)
Money/job	5	3 (6%)
Nonspecific progress	5	2 (4%)
Screen time	4	2 (4%)
Time management	3	3 (6%)
Comparisons to others	2	2 (4%)
Spirituality	2	2 (4%)
Miscellaneous	1	1 (2%)

[a]Percentages exceed 100% because participants could indicate as many metrics as they would like.

TABLE 1 Demographics of the complete sample and by condition

Characteristic		Complete sample ($N = 48$)	Clinician-determined group ($n = 24$)	Clinician + self-determined group ($n = 24$)
Age (M (*SD*))		38.12 (10.86)	38.29 (10.36)	37.54 (10.49)
Gender	Man	23 (48%)	12 (50%)	11 (46%)
	Woman	25 (52%)	12 (50%)	13 (54%)
Race/ethnicity	Asian	5 (10%)	4 (17%)	1 (4%)
	Black or African American	7 (15%)	3 (13%)	4 (17%)
	Hispanic or Latinx	8 (17%)	2 (8%)	6 (25%)
	White	27 (58%)	14 (58%)	13 (54%)
	Prefer to self-identify	1 (2%)	1 (4%)	0 (0%)
Education	High school graduate	3 (6%)	1 (4%)	2 (8%)
	Some college	9 (19%)	3 (13%)	6 (25%)
	College graduate	26 (54%)	13 (54%)	13 (54%)
	Post graduate coursework	10 (21%)	7 (29%)	3 (13%)
Household income	Less than $25,000	1 (2%)	0 (0%)	1 (4%)
	$25,000–$49,999	8 (17%)	3 (13%)	5 (21%)
	$50,000–$74,999	10 (21%)	2 (8%)	8 (33%)
	$75,000–$99,999	10 (21%)	5 (21%)	5 (21%)
	$100,000–$124,999	7 (15%)	6 (25%)	1 (4%)
	$125,000–$149,999	7 (15%)	4 (17%)	3 (13%)
	Over 150,000	5 (10%)	4 (17%)	1 (4%)

TABLE 3 Metrics that participants monitored during the 3 weeks

Metric	Clinician + self-determined group		Clinician-determined group
	n (%[a]) indicated this theme at entry	*n* (%[a]) indicated this theme over time	*n* (%[a]) indicated this theme
Theme: food and drinks			
Food intake	2 (8%)	5 (21%)	10 (42%)
Number of days met calorie goal	1 (4%)	1 (4%)	–
Number of days with solid eating	1 (4%)	–	–
Day of binges episodes	1 (4%)	1 (4%)	–
Water intake	1 (4%)	2 (8%)	1 (4%)
Time of meals	1 (4%)	1 (4%)	1 (4%)
Number of meals per day	1 (4%)	1 (4%)	–
Food consumed during a binge	1 (4%)	2 (8%)	–
Number of takeout food orders	1 (4%)	1 (4%)	1 (4%)
Whether chose soda or water for lunch	–	1 (4%)	–
Number of nonwater beverages	–	1 (4%)	–
Triggers of binge eating	–	1 (4%)	–
Weekday vs. weekend binge episodes	–	1 (4%)	–
Number of binge episodes prevented	–	1 (4%)	–
Calorie intake	–	1 (4%)	6 (25%)
Eating log	–	–	1 (4%)
Theme: physiology and body size			
Heart rate	2 (8%)	2 (8%)	–
Blood pressure	1 (4%)	2 (8%)	–
Body fat percentage	1 (4%)	1 (4%)	1 (4%)
Visceral fat percentage	1 (4%)	–	–
Number of pounds lost	1 (4%)	1 (4%)	–
Number of pounds gained	1 (4%)	1 (4%)	–
Relation between changes in binge eating and weight	1 (4%)	–	–
Menstrual cycle	–	1 (4%)	–
Blood oxygen level	–	1 (4%)	–
Waist measurement	–	1 (4%)	–
Body measurement	–	1 (4%)	–
Theme: physical activity			
Amount of activity	2 (8%)	4 (17%)	–
Types of activity	1 (4%)	2 (8%)	–
Number of days active	1 (4%)	1 (4%)	–
Number of days of exercises	1 (4%)	1 (4%)	–
Steps	1 (4%)	1 (4%)	1 (4%)
Exercises	–	1 (4%)	1 (4%)
Activity level	–	1 (4%)	–
Theme: mental health			
Number of meditation sessions	1 (4%)	1 (4%)	–
Number of days having "good willpower"	1 (4%)	–	–
Emotional well-being	1 (4%)	1 (4%)	–
Emotions	1 (4%)	1 (4%)	–
Feelings after a binge	–	1 (4%)	–

TABLE 3 (Continued)

Metric	Clinician + self-determined group		Clinician-determined group
	n (%[a]) indicated this theme at entry	*n* (%[a]) indicated this theme over time	*n* (%[a]) indicated this theme
Theme: sleep			
Sleep time	1 (4%)	1 (4%)	–
Theme: money/job			
Budget	–	1 (4%)	–

[a]Percentages exceed 100% because participants could indicate as many metrics as they would like.

3.4 | Outcomes of monitoring

Table 4 presents the outcomes of self-monitoring by study condition.

3.4.1 | Primary outcomes

Regarding overall satisfaction, there was no significant difference in satisfaction scores between groups (U = 232.5, $p = .26$). Regarding compliance, participants in the "clinician-determined monitoring" condition submitted 135 entries on time, and 18 participants (75%) submitted at least six entries on time. Participants in the "clinician + self-determined monitoring" condition submitted 137 entries on time, and 21 participants (88%) submitted at least six entries on time. There was no significant difference between study conditions in the proportion of compliant participants ($p = .46$).

3.4.2 | Secondary outcomes

Regarding frequency of monitoring, the highest proportion of participants in both the "clinician-determined monitoring" condition (54%) and the "clinician + self-determined monitoring" condition (46%) indicated that monitoring twice a week was not enough; this was not significantly different ($p = .34$). Most participants in both conditions reported that weight was the most important metric to monitor ("clinician-determined monitoring" condition: 79%; "clinician + self-determined monitoring" condition: 88%). As for likelihood of continuing to monitor without monetary reward, there was no significant difference between groups ($t(46) = 0.07$, $p = .94$).

TABLE 4 Outcomes of monitoring

Outcomes	Clinician-determined group	Clinician + self-determined group
Overall satisfaction (Median (IQR))	5.00 (1.64)	5.86 (1.89)
Compliance		
Total entries on time	135 (94%)	137 (95%)
Total entries	142 (99%)	148 (103%)
Number submitted 6+ entries on time	18 (75%)	21 (88%)
Total people submitted 6+ entries	22 (92%)	23 (96%)
Satisfaction of monitoring frequency		
Not enough	13 (54%)	11 (46%)
Just right	11 (46%)	10 (42%)
Too much	0 (0%)	3 (13%)
Likelihood of continuing to monitor (M (SD))	8.33 (2.39)	8.38 (2.39)
Perception of the most important metrics to monitor[a]		
Weight	19 (79%)	21 (88%)
Number of binge episodes	17 (71%)	19 (79%)
Metric they monitored[b]	NA	13 (54%)
Other metric	9 (38%)	4 (17%)
Nothing	0 (0%)	0 (0%)

[a]Percentages exceed 100% because participants could indicate as many metrics as they would like.
[b]Assessed only among participants in the "clinician + self-determined" group.

4 | DISCUSSION

This study examined the metrics that consumers want to self-monitor in a mobile intervention for binge eating and weight management and tested in a proof-of-concept randomized trial whether monitoring different metrics would lead to differences in satisfaction and compliance with self-monitoring. Results inform the design of the self-monitoring component within a future mobile intervention for binge eating and weight management.

First, findings showed that consumers value a large variety of outcomes besides binge eating and weight. This finding aligns with a recent study on depression showing that substantially more outcomes matter to patients, informal caregivers, and healthcare professionals than the typical outcome domains assessed for depression (Chevance et al., 2020). Identifying the metrics consumers want to monitor is

useful for informing the design of the self-monitoring component of interventions and supports the larger goal of designing interventions that help consumers achieve their goals (Munson et al., 2020). Our results suggest it may be useful for binge eating and weight management interventions to engage consumers in identifying and monitoring additional metrics beyond binge eating and weight that matter to them. Although our study cannot confirm the optimal number of metrics that consumers should monitor (nor that there even is an optimal number of metrics), this work moves us closer toward the goal of designing interventions that meet the needs of its consumers to achieve engagement and improved clinical outcomes.

To that end, we also tested whether encouraging consumers to monitor metrics that matter to them may help with engagement (i.e., satisfaction and compliance). Results of our proof-of-concept trial showed that monitoring a self-determined metric, compared to only monitoring clinician-determined default options, over a brief window of time did not lead to statistically significant differences in consumer satisfaction or compliance. Therefore, our concept was not proven, which might suggest against delivering self-determined monitoring in this population. However, there were nuances of the current study that might have contributed to these null quantitative results and are important to consider. Our study period may have been too short to detect differences. Additionally, there could be a ceiling effect in participants' satisfaction scores.

Another possible explanation is that nearly all participants in the "clinician-determined" condition independently monitored extra metrics beyond the instructed two, which made their monitoring experience similar to participants in the "clinician + self-determined" condition. This was an interesting, unexpected discovery. Our study was designed with two experimental conditions, and—like self-monitoring worksheets or in digital interventions where there typically are pre-set tracking options—each condition was administered the appropriate number of text boxes to enter their specified metrics. However, we learned through our qualitative questions that our manipulation became compromised, as participants' qualitative reflections revealed that they were tracking more metrics on their own than was requested. This is an interesting discovery because it reflects peoples' experiences as they occur in the "real world," but would have been missed using the traditional quantitative self-monitoring options. Therefore, our findings provide insight into what consumers are actually doing in their day-to-day lives, from which we can be better equipped to design monitoring tools that align with consumers' needs and preferences. Consequently, although our tested concept was not proven here, our results do seem to suggest that it may be beneficial to create opportunities for more flexibility for self-monitoring that fit with consumers' own interests within digital interventions.

Indeed, our finding that nearly everyone monitored additional metrics suggests that consumers may value being able to monitor both default and self-determined metrics in a digital intervention; this is consistent with evidence for the importance of identifying individualized goals, and designing monitoring systems to support those goals, for individual health management (Munson et al., 2020). The variety of metrics chosen in our study suggests that an intervention may benefit from a design that offers consumers a customizable option for monitoring, as results did not point to prioritizing one particular metric aside from binge eating and weight, which most participants still perceived as most helpful to monitor. This finding, along with the lack of significant differences in satisfaction and compliance between conditions, supports sustaining standard self-monitoring metrics in mobile interventions for binge eating and weight management. It also parallels findings in other self-tracking domains that people often both want expert (i.e., clinical) guidance on relevant data to track as well as to have the option of tracking data in support of testing their own ideas about contributors and outcomes (Lordon et al., 2020; Schroeder et al., 2019).

When digital interventions include a customizable monitoring option, our findings suggest consumers may also need support to make good decisions about whether, when, and how to switch metrics. Even in this short study, many participants switched the metrics they were monitoring during the monitoring period, despite instructions to monitor the same metrics over the 3 weeks. It could be problematic to change metrics before consumers have enough data to see and learn from trends. In other situations, it may be beneficial to switch, for example, if someone has learned all they can from a metric or have adjusted their behavior, then turning to new metrics may be helpful. Therefore, evaluating designs that support self-monitoring consistency and/or decision rules for changing metrics, over a longer duration, represents a useful area for future research (e.g., Epstein et al., 2021).

4.1 | Strengths and limitations

This study is the first to our knowledge to identify consumers' preferences for self-monitoring in a digital intervention for binge eating and weight management. Furthermore, this study is the first proof-of-concept randomized trial that assessed the impact of different self-monitoring designs on engagement. Finally, this study is strengthened by the use of mixed methods qualitative and quantitative data to enhance our understanding of participants' experiences.

There were also study limitations. One, although our study condition manipulation was designed to create two separate groups, participants in the "clinician-determined" condition monitored additional metrics on their own, compromising the manipulation. Two, participants were paid for completing the study, which may have affected their compliance. Three, the self-monitoring activity was not conducted as part of a behavior change intervention; as a proof-of-concept prototyping activity, we only asked participants to monitor weight and binge eating, but evidence-based treatments can involve monitoring more metrics. Also, unlike in treatment, participants monitored these metrics autonomously without support from a clinician or guide, which may have hindered their understanding of what to do with these metrics and which could have contributed to their switching behaviors. Four, our sample was derived from a registry of individuals who were signed up for an app, dscout, which advertises qualitative market research opportunities. Most participants were at

least college educated, meaning that experiences of dscout users may not transfer to a broader national population despite more balance in other demographics (i.e., race, ethnicity, gender). Additionally, people who sign up for dscout may be more motivated to share in-the-moment insights of their real-time experiences, which may limit generalizability to actual intervention delivery. However, most research studies require participants to be interested in and consent to participate, similar to how our participants needed to be interested in recording and sharing their real-time data and behaviors related to self-monitor to participate. Therefore, although the sample may be more motivated, self-monitoring and sharing real-time data are needed to some extent in real-world interventions with self-monitoring. Five, this study's findings are specific to people with recurrent binge eating and obesity; we cannot assume findings can be generalized to people with other eating disorders/weight statuses without additional study. Finally, the sample size was relatively small, so the quantitative analyses could be underpowered.

5 | CONCLUSION

This study examined consumers' self-monitoring preferences and the relationship among engagement (satisfaction, compliance) and selection of monitoring metrics as applied to a mobile intervention for binge eating and weight management. Our findings suggest that there are a large variety of metrics besides binge eating and weight that matter to consumers, although these two metrics are still of highest value. Our proof-of-concept trial results did not show statistically significant differences in compliance or satisfaction between the two monitoring conditions, indicating our concept under investigation was not proven. However, results suggested design implications of encouraging an individually customizable monitoring option and support for deciding when and how to switch metrics that are being monitored, which need to be tested in future research. Future studies also would benefit from evaluating self-monitoring over a longer period during actual mobile intervention delivery.

ACKNOWLEDGMENT

This study was supported by a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (K01 DK116925).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Jianyi Liu https://orcid.org/0000-0002-5065-1771
Sean A. Munson https://orcid.org/0000-0002-0472-6138
Angela Chang https://orcid.org/0000-0003-3421-5264
Claire Voss https://orcid.org/0000-0003-4652-7476
Andrea K. Graham https://orcid.org/0000-0002-9736-8906

REFERENCES

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2021). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, *21*(1), 37–47. https://doi.org/10.1002/capr.12360

Chevance, A., Ravaud, P., Tomlinson, A., Le Berre, C., Teufer, B., Touboul, S., Fried, E. I., Gartlehner, G., Cipriani, A., & Tran, V. T. (2020). Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: Qualitative content analysis of a large international online survey. *Lancet Psychiatry*, *7*(8), 692–702. https://doi.org/10.1016/s2215-0366(20)30191-7

Chung, C. F., Dew, K., Cole, A., Zia, J., Fogarty, J., Kientz, J. A., & Munson, S. A. (2016). Boundary negotiating artifacts in personal informatics: Patient-provider collaboration with patient-generated data. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, 770–786. https://doi.org/10.1145/2818048.2819926

Chung, C.-F., Gorm, N., Shklovski, I. A., & Munson, S. (2017). Finding the right fit: Understanding health tracking in workplace wellness programs. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 4875–4886. https://doi.org/10.1145/3025453.3025510

Cordeiro, F., Epstein, D. A., Thomaz, E., Bales, E., Jagannathan, A. K., Abowd, G. D., & Fogarty, J. (2015). Barriers and negative nudges: Exploring challenges in food journaling. Proceedings of the 33rd annual ACM conference on human factors in computing systems, 1159–1162. https://doi.org/10.1145/2702123.2702155

Epstein, D. A., Eslambolchilar, P., Kay, J., Meyer, J., & Munson, S. A. (2021). Opportunities and challenges for long-term tracking. In E. Karapanos, J. Gerken, J. Kjeldskov, & M. B. Skov (Eds.), *Advances in longitudinal HCI research* (pp. 177–206). Springer International Publishing. https://doi.org/10.1007/978-3-030-67322-2_9

Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. The Guilford Press.

Folkersma, W., Veerman, V., Ornée, D. A., Oldehinkel, A. J., Alma, M. A., & Bastiaansen, J. A. (2021). Patients' experience of an ecological momentary intervention involving self-monitoring and personalized feedback for depression. *Internet Interventions*, *26*, 100436. https://doi.org/10.1016/j.invent.2021.100436

Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). A tipping point for measurement-based care. *Psychiatric Services*, *68*(2), 179–188. https://doi.org/10.1176/appi.ps.201500439

Gao, M., Kortum, P., & Oswald, F. (2018). Psychometric evaluation of the USE (usefulness, satisfaction, and ease of use) questionnaire for reliability and validity. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1414–1418. https://doi.org/10.1177/1541931218621322

Graham, A. K., Munson, S. A., Reddy, M., Neubert, S. W., Green, E. A., Chang, A., Spring, B., Mohr, D. C., & Wildes, J. E. (2021). Integrating user-centered design and behavioral science to design a Mobile intervention for obesity and binge eating: Mixed methods analysis. *JMIR Formative Research*, *5*(5), e23809. https://doi.org/10.2196/23809

Graham, A. K., Neubert, S. W., Chang, A., Liu, J., Fu, E., Green, E. A., Kornfield, R., & Nicholas, J. (2021). Applying user-centered design methods to understand Users' day-to-day experiences can inform a mobile intervention for binge eating and weight management. *Frontiers in Digital Health*, *3*, 42. https://doi.org/10.3389/fdgth.2021.651749

Graham, A. K., Wildes, J. E., Reddy, M., Munson, S. A., Barr Taylor, C., & Mohr, D. C. (2019). User-centered design for technology-enabled services for eating disorders. *The International Journal of Eating Disorders*, *52*(10), 1095–1107. https://doi.org/10.1002/eat.23130

Karkar, R., Schroeder, J., Epstein, D. A., Pina, L. R., Scofield, J., Fogarty, J., Kientz, J. A., Munson, S. A., Vilardaga, R., & Zia, J. (2017). TummyTrials: A feasibility study of using self-experimentation to detect individualized food triggers. Proceedings of CHI Conference on Human Factors in Computing Systems, 6850–6863. https://doi.org/10.1145/3025453.3025480

Kim, J. P., Sadeh-Sharvit, S., Darcy, A. M., Neri, E., Vierhile, M., Robinson, A., Tregarthen, J., & Lock, J. D. (2018). The utility and acceptability of a self-help smartphone application for eating disorder behaviors. *Journal of Technology in Behavioral Science*, *3*(3), 161–164.

Latner, J. D., & Wilson, G. T. (2002). Self-monitoring and the assessment of binge eating. *Behavior Therapy*, *33*(3), 465–477. https://doi.org/10.1016/S0005-7894(02)80039-9

Lordon, R. J., Mikles, S. P., Kneale, L., Evans, H. L., Munson, S. A., Backonja, U., & Lober, W. B. (2020). How patient-generated health data and patient-reported outcomes affect patient-clinician relationships: A systematic review. *Health Informatics Journal*, *26*(4), 2689–2706. https://doi.org/10.1177/1460458220928184

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface*, *8*(2), 3–6.

Lyon, A. R., & Koerner, K. (2016). User-centered design for psychosocial intervention development and implementation. *Clinical Psychology*, *23*(2), 180–200. https://doi.org/10.1111/cpsp.12154

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & 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

Munson, S. A., Schroeder, J., Karkar, R., Kientz, J. A., Chung, C.-F., & Fogarty, J. (2020). The importance of starting with goals in N-of-1 studies. *Frontiers in Digital Health*, *2*, 3. https://doi.org/10.3389/fdgth.2020.00003

Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.

Norman, D. A., & Draper, S. W. (1986). *User centered system design: New perspectives on human-computer interaction*. CRC Press.

Schroeder, J., Chung, C.-F., Epstein, D. A., Karkar, R., Parsons, A., Murinova, N., Fogarty, J., & Munson, S. A. (2018). Examining self-tracking by people with migraine: Goals, needs, and opportunities in a chronic health condition, Proceedings of the 2018 Designing Interactive Systems Conference, Hong Kong, China. https://doi.org/10.1145/3196709.3196738

Schroeder, J., Karkar, R., Murinova, N., Fogarty, J., & Munson, S. A. (2019). Examining opportunities for goal-directed self-tracking to support chronic condition management. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *3*(4), 151. https://doi.org/10.1145/3369809

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, *78*(3), 298–311. https://doi.org/10.1037/a0019247

Torous, J., & Roux, S. (2017). Patient-driven innovation for mobile mental health technology: Case report of symptom tracking in schizophrenia. *JMIR Mental Health*, *4*(3), e27. https://doi.org/10.2196/mental.7911

Tregarthen, J. P., Lock, J., & Darcy, A. M. (2015). Development of a smartphone application for eating disorder self-monitoring. *The International Journal of Eating Disorders*, *48*(7), 972–982. https://doi.org/10.1002/eat.22386

Tregarthen, J., Paik Kim, J., Sadeh-Sharvit, S., Neri, E., Welch, H., & Lock, J. (2019). Comparing a tailored self-help mobile app with a standard self-monitoring app for the treatment of eating disorder symptoms: Randomized controlled trial. *JMIR Mental Health*, *6*(11), e14972. https://doi.org/10.2196/14972

Weinheimer, E. A., Chang, A., Neubert, S. W., Wildes, J. E., & Graham, A. K. (2020). Past, current, and future willingness to engage with treatment targets: Applying user-centered design to inform the design of a mobile behavioral intervention. *The International Journal of Eating Disorders*, *53*, 611–617. https://doi.org/10.1002/eat.23252

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Liu, J., Munson, S. A., Chang, A., Voss, C., & Graham, A. K. (2022). Understanding self-monitoring to inform a mobile intervention for binge eating and weight management: A proof-of-concept randomized trial. *International Journal of Eating Disorders*, *55*(5), 642–652. https://doi.org/10.1002/eat.23700