Original Research Article

DIGITAL HEALTH

DIGITAL HEALTH
Volume 11: 1–12

DOI: 10.1177/20552076241312569
journals.sagepub.com/home/dhj
Sage

Fluctuations in mHealth engagement following receipt of goal-discrepant feedback messages

Lex Hurley[1], Nisha G O'Shea[2], Brooke T Nezami[3], Carmina G Valle[3,4] and Deborah F Tate[1,3,5]

Abstract

Objective: Digital behavior change interventions can successfully promote change in behavioral outcomes, but often suffer from steep decreases in engagement over time, which hampers their effectiveness. Providing feedback on goal performance is an established technique to promote goal attainment; however, theory indicates that sending goal-discrepant feedback messages could cause some users to respond more negatively than others. This analysis assessed whether goal-discrepant messaging was negatively associated with participant engagement, and if this relationship was exacerbated by baseline depressive symptoms within the context of a three-month weight loss pilot mHealth intervention.

Methods: This analysis applied a generalization of log-linear regression analysis with $n = 52$ participants (78.8% female, 61.5% white, ages 21–35) to assess the likelihood of reading consecutive program messages following receipt of messages with goal-discrepant content.

Results: Receipt of goal-discrepant messages was associated with a significantly lower likelihood (RR = 0.89) of participants reading the next program message sent, compared to receiving positive/neutral messages or no message, but these relationships were not influenced by depressive symptoms in this sample.

Conclusion: Feedback on goal performance remains an important behavior change technique; however, sending push messages that alert participants to their goal-discrepant status seems to reduce the likelihood that participants will read future program messages. Sending messages containing positive or neutral content does not seem to carry this negative risk among individuals in goal-discrepant states.

Keywords

mHealth, eHealth, engagement, depression, weight

Submission date: 8 August 2024; Acceptance date: 22 December 2024

Introduction

Overweight and obesity are major contributors to preventable mortality and morbidity, including certain types of cancers, and can negatively impact physical, mental, and emotional health, with a strong correlation to depression.[1–3] Life adjustments from the COVID-19 pandemic have been associated with increased weight gain across the US, with people affected by anxiety and/or depression exhibiting statistically greater weight gains relative to those without such conditions.[4,5] Additionally, population statistics have shown that the number of people reporting depression and anxiety symptoms have increased >300%

[1]Department of Health Behavior, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[2]Research Triangle Institute International, Research Triangle Park, North Carolina, USA
[3]Department of Nutrition, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[4]Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[5]Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, USA

Corresponding author:
Lex Hurley, Department of Health Behavior, University of North Carolina at Chapel Hill, 1700 MLK Jr. Blvd, Chapel Hill, NC 27514, USA.
Email: lexh@unc.edu

from 2019–2021 during the pandemic.[6] For these reasons, it is likely that living through stay-at-home isolation for pandemic safety and later adjusting to work-from-home status in the following years has increased the need and demand for mobile health (mHealth) and wellness programs.

Digital behavior change interventions (DBCIs), including eHealth and mHealth approaches, have been used for decades as effective, scalable alternatives to traditional in-person interventions, and have been repeatedly shown to be more effective than control or usual care comparisons.[7–11] In modern DBCIs, the information, tools, and resources to promote behavior change are typically located inside a website or app, which can be tracked to monitor a participant's engagement with the digital interface, and indicate the dose received of intervention contents. In their critical interpretive synthesis of engagement with DBCIs, Perski et al. defined engagement as "(1) the extent (e.g. amount, frequency, duration, and depth) of usage and (2) a subjective experience characterized by attention, interest, and affect," which is applied here.[12] It is well known that DBCIs often experience sharp declines in participant engagement over time, as participants may open and use the program less over time, or disengage entirely.[13] Some research has indicated that mental health issues including depression are likely associated with reduced engagement, which could exacerbate this trend.[12,14] This limits their opportunity to interact with intervention contents designed to improve their behaviors and likely reduces the overall success of the intervention. As such, it is helpful to understand how to keep participants engaged with DBCIs.[15]

Many DBCIs apply behavior change techniques (BCTs) related to feedback on goal status, whereby a user may compare their logged performance with their current goals or program recommendations and, in some programs, receive suggestions to help attain goal success and behavior change.[16–19] Traditional interventions usually involve in-person program counselors who can read non-verbal cues, empathize with participants, and adapt their phrasing or delivery of their feedback and suggestions to keep sessions constructive and avoid distressing participants. One shortcoming of DBCIs is that they are often unable to read, anticipate, or tailor participant reactions to program communications or components.[20–22]

Messaging in digital programs is often pre-written by researchers and study staff according to predetermined tailoring variables (e.g. gender, age, etc.) and decision rules based on program performance (e.g. adherence to dietary goals, weight change, etc.), and can include a variety of topics including praise, reinforcement, reminders, and feedback on behavioral or outcome goals.[23–29] DBCIs can deliver this feedback in various formats, including graphs and icons which may passively communicate goal progress, via short push messages designed to get users' attention, or via longer weekly feedback summary messages. Of particular interest in this analysis are push feedback messages that highlight the discrepancy between one's current behavior and a goal, with the intent to cue a participant to focus attention on reducing this discrepancy and achieving the goal (BCT 1.6).[19] The actual effectiveness of these types of messages in DBCIs is not well studied and may have heterogeneous impacts across participants based on some theoretical determinants.[16,17,30]

Theoretical foundations

There is some phenomenological overlap between describing a tendency for some individuals to derive negative interpretations of incoming information, which can cause strong negative affective/emotional reactions, and can negatively influence future behaviors.[31–33] The Cognitive Theory of Depression describes "negative information processing biases" as a tendency for individuals to selectively discount or ignore positive feedback and focus on negative feedback.[31] Affected individuals tend to have stronger reactions to negative information and have shown greater recall of negative information relative to positive, which can play a role in shaping their worldviews and expectations, greatly increasing their risk of developing depression.[31] This is related to an expansion of Attribution Theory referred to as having "pessimistic attributional styles" which can carry meaningful influence in behavior change interventions.[32]

Briefly, Attribution Theory states that when confronted with any type of achievement information, people will attempt to causally attribute their reason(s) for success or failure to a locus, stability, and controllability of the cause as "naïve psychologists," which may then influence their emotional, affective, cognitive, and behavioral reactions.[32–34] While these attributions may theoretically occur following success or failure, they tend to exhibit stronger, lasting impacts following exposure to goal-discrepant feedback, so this will be the primary focus in this study.[32] It is cognitively taxing to uniquely attribute cause(s) for each event, therefore, individuals tend to develop their own attributional styles as a heuristic likely explanation of events.[35] Those with "pessimistic attributional styles" have a tendency to attribute failures to internal, stable, and uncontrollable factors—a failing within themselves that will persist and that they cannot change—which can cause strong negative reactions including shame, helplessness, anxiety for future goal performance, and possibly compromise their self-efficacy to achieve that goal in the future.[32,35–40]

These tendencies are not exclusive to individuals with severe depression, but rather confer vulnerabilities to developing depression, and could be common among participants recruited for DBCIs.[31,38] Hypothetically, if a DBCI pushed a goal-discrepant message to participants with these traits, they would likely experience a strong negative emotional reaction. With no way to observe this issue and rectify messaging patterns, the DBCI will likely send more of these types of messages over time, which could lead to further deleterious effects. For example, participants may associate future messages as having a risk of containing this type of information and avoid reading them, lowering their program engagement, which could then contribute to reduced changes in health outcomes.[41,42]

The current study tests these theoretical pathways via the following hypotheses: (1) Participants will be less likely to view the next program message sent after viewing a message containing goal-discrepant content, compared to messages with other types of content or no message, within the context of a mHealth weight management intervention. (2) This negative relationship will be moderated by depressive symptoms such that those with higher baseline depressive symptoms will be less likely to view consecutive messages after viewing those containing goal-discrepant content compared to those with lower baseline depressive symptoms.

Methods

Study design and participants

Data for this secondary analysis come from the Nudge pilot study, a 12-week just-in-time adaptive intervention (JITAI) studying the effects of microrandomized intervention messages on the achievement of daily behavioral goals to promote weight loss (clinicaltrials.gov identifier NCT03836391). The study was approved by the University of North Carolina Institutional Review Board (#16-0775). The IRB waived written documentation of informed consent and consent was obtained via an online consent form with electronic agreement prior to enrollment and data collection. Nudge recruited 53 young adults aged 18–35 living with overweight or obesity (BMI between 25 and 40 kg/m^2), who reported < 150 min of weekly moderate-to-vigorous physical activity, owned an iPhone (Apple, Cupertino, CA, USA), had not been pregnant within the past 6 months. One participant became pregnant prior to completing the study and withdrew, rendering the effective sample size for this analysis $n = 52$.

Participants received a wearable activity tracker and wireless scale (Fitbit, San Francisco, CA, USA), and downloaded the Nudge study app containing lessons, resources, self-monitoring tools, personalized daily goals for diet, exercise, and daily weighing, as well as tailored weekly feedback and daily tailored messages targeting seven BCTs. Each participant had tailored daily goals for weighing, dietary consumption, and physical activity. An example screenshot of the app is shown in Figure 1. Data collection occurred from February–September 2019. For further details on the Nudge pilot intervention, see Valle, Nezami, Tate (2020).[43]

Nudge was a microrandomized trial (MRT), which is an mHealth intervention design that enables empirical

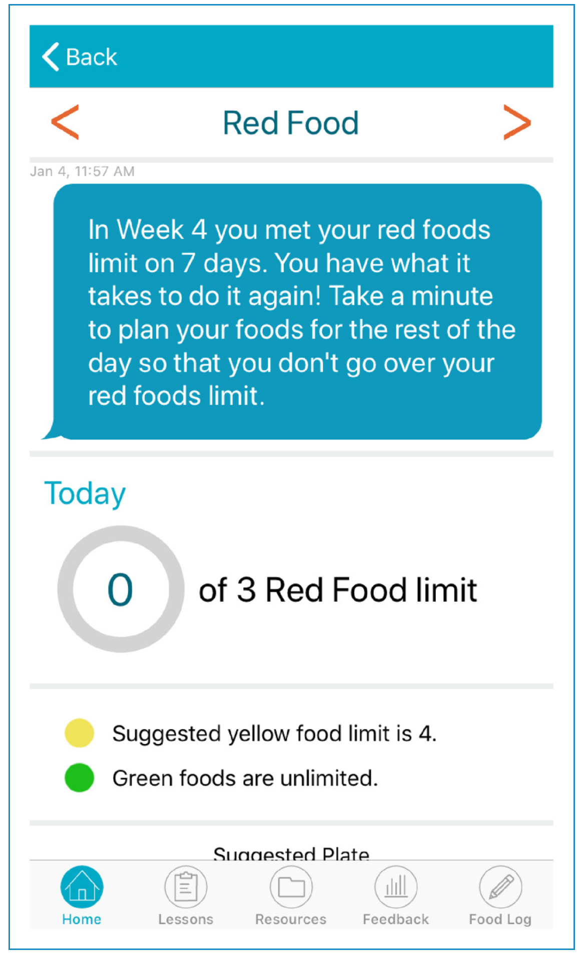

Figure 1. Screenshot from Nudge app.

assessment of message impact within a DBCI by randomizing the probability of any message being sent according to a known value and collecting covariate data at each decision point to measure if a user's performance changed during times when they did or did not receive a message, and can be used to inform the development of JITAIs.[43–47] Nudge applied 4 daily decision points (early morning = 7:00AM, late morning = 10:00–12:00PM, afternoon = 2:00–4:00PM, evening = 7:00–9:00PM) when its algorithm would assess participant eligibility for different message types based on program decision rules, select one of these messages at random, then randomly determine whether to send the message with a 0.5 probability of sending or null.[43] Messages would remain visible on the app until midnight each program day, then expire. Each message focused on a single behavior, and program decision rules ensured no more than one message was delivered for each of the three goal behaviors (dietary goals/logging, physical activity, daily weighing) would be delivered in a single day. Covariate data were collected at each of the four daily decision points, resulting in $N = 16{,}425$ total observation points clustered across 4368 person days, permitting detailed analysis of time-varying proximal engagement outcomes. To control for confounding in this analysis, all observations are subset to decision points when users were available to be messaged and in a goal-discrepant state, defined as not meeting ≥1 behavioral goal at that decision point. This reduced the number of available observations to $n = 12{,}920$.

Measures

Viewing the next pushed message sent is the binary dependent variable (DV) for all analyses. Message viewed indicators were available at all decision points, which were used to calculate the time-varying DV as follows: For each decision point at time t (whether a message was sent or not), code 1 if the next message sent at time $t+1$ was viewed before the following message sent at time $t+2$ was sent, and 0 if not. These $t+1$ gaps could vary in length if users were sent consecutive decision point nulls; in this case, each of the consecutive nulled points would become 0 or 1 based on response to the next message sent. Additionally, due to the nature of relying on the next message viewed as the dependent variable, it was necessary to drop the last message(s) sent for each user as NAs, which further reduced the number of eligible observations from 12,920 to a final effective $n = 12{,}774$. This variable serves as a manifest indicator for the latent construct of proximal program engagement, which has been acknowledged and validated in previous studies.[12,48,49]

Types of message content. Nudge message libraries were dummy coded to identify messages containing goal-discrepant feedback. Certain types of messages could only be sent to users when they were in goal-discrepant states according to program decision rules and were coded as such. Other message types such as social comparison for participant behavioral goals (e.g. weighing, active minutes, and dietary tracking) were always eligible to be sent. These social comparison messages could indirectly convey goal-discrepant information if they were sent while a participant was in a goal-discrepant state for the corresponding behavior (i.e., a participant who has not currently achieved their active minutes goal receives a message stating "69% of Nudge participants have achieved their active minutes goals so far today! Are you one of them?"). As such, social comparison messages were cross-referenced between the time they were sent and if the participant had currently achieved that behavioral goal at the time of receipt. If they had not achieved that goal, the message was dummy-coded as containing goal-discrepant information. Additionally, outcome feedback messages specifying that a participant had gained >1 lb of weight since the last time they weighed were coded as goal-discrepant, as this was a weight management intervention.

The remaining sent messages were then categorized as containing positive/neutral content. Multiple independent variables (IVs) were then created based on these dummy codes for contrast analyses (detailed below). The first author (LH) conducted sensitivity checks by visually inspecting message content relative to dummy codes and manually making adjustments where necessary to ensure accuracy.

Depressive symptoms. Depressive symptoms are indicated by baseline scores on the Centers for Epidemiology Scale – Depression (CES-D), collected as part of an online baseline survey, as an effect indicator for depressogenic attributions. As the Cognitive Theory of Depression posits that depressogenic attributions increase the likelihood of developing depression, it is reasonable that those exhibiting higher symptoms are also likely to have higher depressogenic attributions, meeting criteria for an effect indicator influenced by the latent construct of interest.[31,50]

Covariates. Age, gender, and race/ethnicity were applied as time-invariant sociodemographic covariates, as it is known that the CES-D is biased to represent white, female manifestations of depressive symptoms, which could influence response validity.[50,51] Program week was applied as a time-varying covariate used to control for the general passage of time and reduce noise in model estimations based on work by Murphy & Almirall.[52]

Statistical analysis

Analyses follow the weighting and centering method for analyzing causal time-varying effects in the presence of time-varying confounders developed by Boruvka et al. and expanded by Qian et al.[53,54] To the authors' knowledge, there is currently not an established power calculation method for this type of analysis.[54] Borrowing model notation from Boruvka et al., a user's treatment values are denoted by A_t and precede their subsequent proximal response Y_{t+1}.[53] Users are considered available ($I_t = 1$) if they were eligible to receive a treatment message at a given decision point. All observations in this analysis are subset to be considered available. User covariates and potential moderators are represented by the vector $\boldsymbol{X_t}$. When arranged in longitudinal order over T treatment occasions, these variables follow the pattern (X_T, A_T, Y_{T+1}). A user's history $\boldsymbol{H_t}$ is a vector consisting of all previous information accrued up until a given treatment at time t, $H_t = (\bar{X}_t, \bar{Y}_t, \bar{A}_{t-1})$. The general formula representing the effects of treatment on the proximal outcomes is: $E[Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0) \mid S_{1t}(\bar{A}_{t-1})]$, where $S_{1t}(\bar{A}_{t-1})$ is a subset of the user's full history H_t. Conditioning on only a subset of the full user history enables marginalized effect size observations to test different hypotheses using the same data.[53] Confounders included in the weights help to improve the plausibility that the coefficient A_t reflects a causal effect, as it contrasts the conditional mean between the proximal responses if a user who was eligible to receive a goal-discrepant message was randomized to receive ($a_t = 1$) or not receive ($a_t = 0$) a message during the program. This method is based on Robins's previous work for a marginal generalization of the treatment "blips," or temporal deviations from one's expected activity given their longitudinal trajectory within structural nested mean models, conditioning on a subset of hypothesized variables of interest, then weighting and centering longitudinal treatment effects before conducting least-square estimation for treatment effect size.[53,55] Taken together, this method calculates the estimated marginal excursion effect (EMEE, represented by $\hat{\beta}_0$) of treatment to determine if significant deviations in observed response vary from each user's expected patterns from times they received treatment based on the IV, compared to the times they did not.

Stated plainly, this method essentially generates an expected trajectory pattern of usage for each participant across all of their observations over time. At each observation point (t), the method calculates a participant's expected value at $t + 1$ if they were to have received, or not received, the message treatment at time t, then compares this to the value actually observed at time $t + 1$. The EMEE is then calculated as an overall deviation value of these [observed – expected] differences across all participants over time, conditioned on covariates; representing an overall "shock" to a participant's trajectory after viewing a message at time t. For this reason, these models control for whether messages sent at time t were ignored, as participants who ignore messages over a length of time would by definition show no difference from expected values based on their trajectories at that time, based on IVs. Therefore, any significant deviations from their expected value, indicated by the EMEE, are likely to result from having viewed a given message at time t.

These models can test for effect moderation ($\hat{\beta}_1$), that is, whether measures of the EMEE significantly vary according to a given moderator variable. These models are only capable of examining a single moderation effect at a time for a dataframe conditional on a subset of variables relevant to study hypotheses. Estimation typically assumes a linear model for treatment effects; however, this study incorporates a logit link function to accommodate binary outcomes using R code Qian et al. has posted to Github repository.[54,56] All analyses were completed using R statistical software.[57]

Model building

Model 1 tests hypothesis 1 by measuring the EMEE of sending any message compared to sending no message at a decision point ($\hat{\beta}_0$). It applies dummy-coded indicators for whether a sent message was coded as containing goal-discrepant information as the effect moderator ($\hat{\beta}_1$), and includes sex_i, race_i, mean-centered age_i, standardized CES-D scores_i, and program week_{it} as control variables (S_{itT}).

Models 2–4 test hypothesis 2 for moderation from CES-D scores. To fully probe the potential moderating influence of depressive symptoms, it was necessary to create a series of contrast models measuring the EMEE of different types of message exposures using IVs with different dummy-codes: Model 2 compares goal-discrepant messages vs. null, Model 3 compares positive/neutral messages vs. null, and Model 4 compares goal-discrepant messages vs. positive/neutral messages; all of which use baseline CES-D scores as the effect moderator ($\hat{\beta}_1$) and the same control covariates. The Nudge dataframe was necessarily subset according to both the relevant contrast IVs and non-missing values of the next message viewed DV, which led to variability in the effective sample size in each model and are noted in the titles of each table.

Results

Descriptive statistics

Participants were on average 29.5 (SD = 3.8) years old, with an average body mass index (BMI) of 31.9 kg/m^2 (SD = 4.3). 78.8% of participants self-identified as female, and 38.5% self-identified as having racial/ethnic minority backgrounds. The sample was highly educated, as most participants (84.9%) had a college degree. Additionally, the sample reported low depressive symptoms overall, with an average baseline CES-D score of 9.27 (SD = 7.4), median score of 6, and only10 participants (19.2%) meeting the cut point of ≥16 to indicate risk of clinical depression.[58] Full baseline demographic characteristics are displayed in Table 1.

Out of the $n = 12{,}774$ subset decision points for this secondary analysis, 6210 messages were sent to participants. This amounted to approximately 119.4 messages delivered to each participant on average (SD = 19.31), with a median of 120.5 messages, and a range of 49:154 messages sent in this sample. Among these, approximately 79.44 messages were read (~67.55%, SD = 28.67), with a median of 84.5 messages read, and range of 9:135 messages read. Of the total 6210 messages sent in the full Nudge sample, approximately 2621 of them (42.2%) were identified as goal-discrepant according to decision rules and visual inspection for this analysis.

Goal-discrepant message effects on proximal engagement

Model 1 measures the EMEE of sending any type of message compared to sending no message at all (β_0), and tests effect moderation of sent messages with the goal-discrepant dummy code, summarized in Table 2. Adjusted standard errors and 95% confidence intervals for small sample size are also presented. All tables report results in the log

Table 1. Participant characteristics of Nudge sample ($n = 52$).

Variable	n (%)	Mean (SD)	Min, Max
Gender			
Male	11 (21.2%)		
Female	41 (78.8%)		
Race			
White	32 (61.5%)		
Black	9 (17.3%)		
Asian	3 (5.8%)		
Other POC, mixed	8 (15.4%)		
Highest education			
High school/GED	2 (3.8%)		
Some college	6 (11.5%)		
College graduate	25 (48.1%)		
Postgraduate degree	19 (36.5%)		
Age (years)		29.5 (3.8)	21, 35
BMI at baseline (kg/m^2)		31.9 (4.3)	25.10, 39.93
BMI at 3-months (kg/m^2)		30.89 (4.4)	22.94, 39.11
CES-D score		9.27 (7.4)	0, 35

SD: standard deviation; POC: people of color; BMI: body mass index; CES-D: Centers for Epidemiology Scale – Depression; 3-month BMI includes only $n = 51$ participants with complete values.

relative-risk scale, with exponentiated relative risks detailed between sections to improve interpretability.

Overall, there was no significant effect on viewing the next message sent after receiving any type of message, meaning that users in goal-discrepant states are approximately just as likely to view the next message a DBCI sends regardless of if they have already received a previous message or no message (MRT nulled) at a given decision point. However, the effect moderation of goal-discrepant content was statistically significant according to normal and adjusted 95% confidence intervals, such that if a user in a goal-discrepant state receives a goal-discrepant push message at time 1, they are approximately 0.89 times less likely to view the next message sent, compared to if they received no message according to this model. For comprehensiveness, this model was also re-run to include all possible user observations regardless of goal-discrepant states and returned similar results.

Model 2 measures the EMEE of sending a goal-discrepant message compared to no message (β_0) and tests effect moderation of CES-D scores, summarized in Table 3.

In agreement with Model 1, pushing goal-discrepant content caused users to be 0.903 times less likely to read the next message sent compared to sending no message, which is statistically significant according to normal and adjusted 95% confidence intervals, but this relationship was not influenced by CES-D scores.

Model 3 measures the EMEE of sending a positive or neutral message compared to no message (β_0) and tests effect moderation of CES-D scores, summarized in Table 4.

According to this model, pushing other types of positive/neutral message content did not influence users' likelihood of reading the next message sent compared to sending no message at all, and this relationship was also not influenced by CES-D scores in this sample.

Table 2. Model 1: Any message versus null; effect moderation of goal-discrepant content ($n = 12{,}545$).

$\hat{\beta}$	Coefficient	SE	Adj. SE	95% CI	Adj. 95% CI
Intercept (β_0)	−0.006	0.018	0.020	−0.042; 0.030	−0.046; 0.033
Goal-discrepant (β_1)	−0.116	0.031	0.035	−0.177; −0.055	−0.187; −0.045
$\hat{\alpha}$					
Intercept	−0.124	0.069	0.079	−0.259; 0.011	−0.278; 0.030
Sex	0.002	0.114	0.132	−0.221; 0.225	−0.257; 0.261
Race	0.016	0.059	0.068	−0.099; 0.132	−0.117; 0.149
Age-mc	−0.009	0.016	0.019	−0.040; 0.022	−0.046; 0.028
Program week	−0.045	0.008	0.008	−0.061; −0.029	−0.060; −0.029
CESD-std	0.054	0.065	0.097	−0.073; 0.181	−0.136; 0.244

Model dimensions: $p = 2$, $q = 6$.
mc: mean-centered; std: standardized; SE: standard error; Adj.: adjusted for small sample size; 95% CI = 95% confidence interval; results displayed in log relative-risk scale.

Table 3. Model 2: Goal-discrepant message versus null; effect moderation of CES-D scores ($n = 5728$).

$\hat{\beta}$	Coefficient	SE	Adj SE	95% CI	Adj. 95% CI
Intercept (β_0)	−0.102	0.022	0.025	−0.146; −0.058	−0.152; −0.052
CES-D (β_1)	−0.027	0.022	0.038	−0.068; 0.015	−0.103; 0.050
$\hat{\alpha}$					
Intercept	−0.106	0.066	0.076	−0.235; 0.023	−0.255; 0.043
Sex	0.044	0.108	0.127	−0.167; 0.255	−0.205; 0.293
Race	0.006	0.060	0.070	−0.111; 0.123	−0.131; 0.143
Age-mc	−0.012	0.015	0.018	−0.041; 0.017	−0.047; 0.023
Program week	−0.042	0.008	0.008	−0.058; −0.026	−0.058; −0.026
CESD-std	0.043	0.064	0.092	−0.082; 0.168	−0.137; 0.223

Model 2 dimensions: $p = 2$, $q = 6$.
mc: mean-centered; std: standardized; SE: standard error; Adj.: adjusted for small sample size; 95% CI: 95% confidence interval; results displayed in log relative-risk scale.

Model 4 measures the EMEE of sending a goal-discrepant message compared to sending a positive or neutral message (β_0), and tests effect moderation of CES-D scores, summarized in Table 5.

According to this model, sending a goal-discrepant message caused users to be 0.892 times less likely to read the next message sent, compared to times when they were sent positive or neutral messages, which is statistically significant according to normal and adjusted 95% confidence intervals, and was not influenced by CES-D scores.

All model estimates support hypothesis 1 that sending a goal-discrepant message caused users to be approximately 0.89 times less likely to read the next message sent relative

Table 4. Model 3: Positive/neutral message versus null; effect moderation of CES-D scores ($n = 8838$).

$\hat{\beta}$	Coefficient	SE	Adj. SE	95% CI	Adj. 95% CI
Intercept (β_0)	−0.004	0.015	0.016	−0.034; 0.025	−0.037; 0.029
CES-D (β_1)	0.019	0.014	0.018	−0.008; 0.047	−0.017; 0.056
$\hat{\alpha}$					
Intercept	−0.047	0.053	0.057	−0.151; 0.057	−0.159; 0.065
Sex	−0.003	0.098	0.113	−0.195; 0.189	−0.224; 0.218
Race	−0.001	0.049	0.055	−0.097; 0.095	−0.109; 0.107
Age-mc	−0.007	0.013	0.015	−0.032; 0.018	−0.301; 0.287
Program week	−0.046	0.007	0.008	−0.056; −0.032	−0.062; −0.030
CESD-std	0.049	0.050	0.062	−0.049; 0.147	−0.072; 0.170

Model dimensions: $p = 2$, $q = 6$.
mc: mean-centered; std: standardized; SE: standard error; Adj.: adjusted for small sample size; 95% CI = 95% confidence interval; results displayed in log relative-risk scale.

Table 5. Model 4: Goal-discrepant message versus positive/neutral message; effect moderation of CES-D scores ($n = 7932$).

$\hat{\beta}$	Coefficient	SE	Adj. SE	95% CI	Adj. 95% CI
Intercept (β_0)	−0.114	0.027	0.032	−0.167; −0.060	−0.179; −0.048
CES-D (β_1)	−0.050	0.034	0.073	−0.116; 0.017	−0.197; 0.097
$\hat{\alpha}$					
Intercept	−0.078	0.060	0.069	−0.196; 0.040	−0.213; 0.057
Sex	0.009	0.106	0.124	−0.199; 0.217	−0.234; 0.252
Race	0.012	0.056	0.064	−0.098; 0.122	−0.113; 0.137
Age-mc	−0.009	0.014	0.018	−0.036; 0.018	−0.044; 0.026
Program week	−0.043	0.007	0.007	−0.057; −0.029	−0.057; −0.029
CESD-std	0.064	0.055	0.075	−0.044; 0.172	−0.083; 0.211

Model dimensions: $p = 2$, $q = 6$.
mc: mean-centered; std: standardized; SE: standard error; Adj.: adjusted for small sample size; 95% CI: 95% confidence interval; results displayed in log relative-risk scale.

to sending no message or sending a positive or neutral message, which is statistically significant. Likewise, sending a positive or neutral message did not seem to make users significantly more or less likely to read future messages compared to sending no message. However, there was no evidence to support the moderation hypothesis 2 that these relationships varied based on CES-D scores in this sample. Considering covariates across all models, only program week exhibited a significant negative effect on the likelihood of reading future messages, which is not unexpected in mHealth programs.

Discussion

This study shows that sending goal-discrepant message content negatively influences proximal engagement outcomes (i.e. viewing subsequent messages) compared to sending other types of messages, or no message. This provides some evidence to warrant caution against including messages with goal-discrepant content reinforcing that a participant has not met their goal in DBCI tailoring algorithms, lest the program contributes to the already decreasing likelihood of engagement over time.

Analyses of MRTs tend to focus on the effects of different types of messages on behavioral outcomes, and are instrumental in informing the design of JITAIs[59] However MRTs are not often studied to assess time-varying impacts on participants' proximal engagement patterns, which could then contribute to downstream changes in proximal and distal behavioral outcomes.[12,15] Research by Bidargaddi et al. using similar methods found that pushing positive- or neutral-toned messages can lead to increased proximal app engagement with users more likely to open an mHealth app following receipt of a push message.[47] However, to our knowledge, this is the first study quantitatively examining message factors that impact the likelihood of reading future messages for proximal engagement within an ongoing DBCI. From a qualitative perspective, Lyzwinski et al. note in a systematic review of consumer perspective studies, some participants noted concern that negative or insensitive messages could create guilt or fear of them failing in the DBCI when not meeting their goals.[60]

Delivering feedback on goal performance remains an important BCT and aspect of goal-setting progress; however, the manner in which this information is delivered must be considered.[18,61] For example, studies informing Goal Setting Theory have shown that individuals with challenging goals tend to maintain or increase performance when issued positive feedback in terms of progress toward a goal (e.g. 75%), but this performance can quickly deteriorate when issued goal-discrepant feedback in terms of disparities against the goal (e.g. −25%).[62,63] Bandura and Locke argue that such feedback can erode one's efficacy beliefs regarding future attainability of the goal and negatively affect their perceived self-efficacy.[30] Likewise this analysis indicates that pushing messages to users reiterating that their goals are not yet achieved before they are due may cause users to be put off from the program and not check back in soon afterwards, and based on these studies possibly risk reducing goal performance overall, as DBCI engagement is associated with improved outcomes.[12] DBCIs in particular must balance the challenge of providing an effective behavior change program in a format that users *want* to return to multiple times per day for months or longer, that can be assimilated into their current or changing lifestyles, and is perceived to be better than some publicly available alternative app which may have a weaker evidence base to support it, or no app at all.[13,60]

Limitations & strengths

This study is not without its limitations. First, Nudge was a pilot intervention that recruited a small number of participants and was not necessarily powered for this type of analysis. As mentioned previously, power calculations are established for this analytical method at the time of this writing,[54] so results should be interpreted with some caution, and replication using larger and more generalizable samples would be beneficial. The CES-D is not an ideal indicator for pessimistic attributional styles, which are more theoretically-aligned predictors than general depressive symptoms, but it was the best proxy indicator given the nature of data available for secondary analysis.[50] The finding that CES-D scores did not moderate the effects goal-discrepant messages exerted on the likelihood of viewing subsequent messages was not unexpected, as the Nudge study had a small number of participants who also reported a fairly low mean CES-D score of 9.27, well below the recognized depression indicator of ≥16, combined with its imperfect function as a proxy indicator. As existing research supports the notion that higher depressive symptoms are associated with lower program engagement, this factor is still likely worth examining in future studies with larger samples and additional scales.[12,38] Additionally, this study is unable to fully control for unmeasured time-varying contextual factors that may potentially influence the observed relationships. This analysis was designed on the premise that the influence of such factors across all participants and observations would wash out into high variance which would make detection of significant effects more difficult but still possible. For example, one potential confounder could be if participants knew they were doing poorly, or knew meeting their goals was not a priority that day, they would avoid reading any messages sent, and thus have a lower likelihood of reading next messages—regardless of message content.[42,64] However, this also likely would have contributed to an overall effect from sending any type of message, which was not observed.

Strengths of the study include the micro-randomized nature of message delivery in Nudge, which helps bolster the plausibility of causality, as participants could not expect or habituate to different types or topics of messages during the intervention. Additionally, the advanced analytical methods shared by Boruvka et al. and Qian et al. provide unique opportunities to measure effects between time-varying IVs, moderators, and DVs while promoting strong evidence for causal arguments provided their method's assurance of association, temporality, and non-spuriousness provided a high number of observations.[53,54,56]

Conclusion

DBCIs seem to excel at motivating successful participants to continue but struggle with how to address waning interest from users who may be struggling with the program. This study suggests that sending messages with positive or neutral content to participants struggling to meet program goals may be preferable to sending messages with goal-discrepant content, which could contribute to lessening program engagement and potentially contribute to risk of disengagement if participants experience strong negative reactions to these types of messages. This will hopefully be helpful for researchers developing tailoring algorithms for upcoming DBCIs. It will be beneficial for future research to examine if these relationships are reproducible among larger groups in longer-duration DBCIs, as well as if these relationships may be predicted by variables measurable at baseline, so as to enable in-depth customization of tailoring algorithms best suited to different participants' preferred communication styles. As a final note, this analysis examines only one potential factor informed by theory which could be contributing to observed reductions in participant engagement, and there may be many other predictable factors contributing to this phenomenon. Rather than suggesting problems from end-users contributing to their lower engagement, it can be beneficial for the field to examine if there are factors in our own programs that may be inadvertently pushing participants away.

Acknowledgements: We would like to thank Matthew Jansen for his assistance with statistical coding to facilitate this analysis.

Contributorship: LH: authored the first draft of the manuscript and implemented co-authors' feedback, conceived the secondary study design, conducted analyses. LH, NGO: methodology. NGO, BTN, CGV DFT: manuscript review. BTN, CGV DFT: main trial design, funding, and data acquisition. All authors read and approved the final version of this manuscript.

Data availability: The authors do not have permission to share raw study data due to requirements to protect the privacy of participants, in accordance with their informed consent. However, de-identified data used in this specific analysis and/or output files related to this analysis will be made available upon request to the corresponding author.

Declaration of Conflicting Interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval: The parent Nudge study was approved by the University of North Carolina Institutional Review Board, #16-0775.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a Cancer Health Disparities training grant T32CA128582 from the National Cancer Institute (NCI) of the National Institutes of Health (NIH). The parent study was supported by a Gillings Innovation Laboratory award and developed in partnership with UNC CHAI Core, which is supported by NIH grant P30DK056350 to the UNC Nutrition Obesity Research Center and NCI grant P30-CA16086 to the UNC Lineberger Comprehensive Cancer Center. The content of this research is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Guarantor: LH

Informed Consent: Informed consent was collected from participants prior to enrollment and data collection. Research staff guided participants through online consent forms via telephone which interested participants completed and submitted through secure online portals.

ORCID iDs: Lex Hurley https://orcid.org/0000-0003-2217-0492
Carmina G Valle https://orcid.org/0000-0003-1985-2409

References

1. Statistics NCfH. Obesity and Overweight (2021, accessed 3/11 2021).
2. Luppino FS, de Wit LM, Bouvy PF, et al. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry* 2010; 67: 220–229.
3. Bianchini F, Kaaks R and Vainio H. Overweight, obesity, and cancer risk. *Lancet Oncol* 2002; 3: 565–574.
4. Khubchandani J, Price JH, Sharma S, et al. COVID-19 pandemic and weight gain in American adults: a nationwide population-based study. *Diabetes Metab Syndr* 2022; 16: 102392.. 20220110.
5. Bhutani S, vanDellen MR and Cooper JA. Longitudinal weight gain and related risk behaviors during the COVID-19 pandemic in adults in the US. *Nutrients* 2021; 13: 671.
6. DeAngelis T. Depression and anxiety escalate during COVID. https://www.apa.org/monitor/2021/11/numbers-depression-anxiety (2021, accessed 5/25 2022).
7. Rivera J, McPherson A, Hamilton J, et al. Mobile apps for weight management: a scoping review. *JMIR Mhealth Uhealth* 2016; 4: e87.
8. Mangieri CW, Johnson RJ, Sweeney LB, et al. Mobile health applications enhance weight loss efficacy following bariatric surgery. *Obes Res Clin Pract* 2019; 13: 176–179.
9. Shaw RJ, Bosworth HB, Silva SS, et al. Mobile health messages help sustain recent weight loss. *Am J Med* 2013; 126: 1002–1009.
10. Flores Mateo G, Granado-Font E, Ferré-Grau C, et al. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. *J Med Internet Res* 2015; 17: e253.
11. Vandelanotte C, Müller AM, Short CE, et al. Past, present, and future of eHealth and mHealth research to improve physical activity and dietary behaviors. *J Nutr Educ Behav* 2016; 48: 219–228.e1.
12. Perski O, Blandford A, West R, et al. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017; 7: 254–267.. 2016/12/15.

13. Eysenbach G. The law of attrition. *J Med Internet Res* 2005; 7: 11.
14. Yeager CM and Benight CC. If we build it, will they come? Issues of engagement with digital health interventions for trauma recovery. *Mhealth* 2018; 4: 37–37.
15. Lin P-H, Grambow S, Intille S, et al. The association between engagement and weight loss through personal coaching and cell phone interventions in young adults: randomized controlled trial. *JMIR Mhealth Uhealth* 2018; 6: e10471.. Original Paper 18.10.2018.
16. Bohlen LC, Michie S, de Bruin M, et al. Do combinations of behavior change techniques that occur frequently in interventions reflect underlying theory? *Ann Behav Med* 2020; 54: 827–842.
17. Carey RN, Connell LE, Johnston M, et al. Behavior change techniques and their mechanisms of action: a synthesis of links described in published intervention literature. *Ann Behav Med* 2018; 53: 693–707.
18. Michie S, Abraham C, Whittington C, et al. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009; 28: 690–701.
19. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013; 46: 81–95.. 2013/03/21.
20. Gilmartin J and Murphy M. The effects of contemporary behavioural weight loss maintenance interventions for long term weight loss: a systematic review. *J Res Nurs* 2015; 20: 481–496.
21. Sorgente A, Pietrabissa G, Manzoni GM, et al. Web-Based interventions for weight loss or weight loss maintenance in overweight and obese people: a. Systematic review of systematic reviews. *J Med Internet Res* 2017; 19: e229.
22. Neve M, Morgan PJ, Jones PR, et al. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. *Obes Rev* 2010; 11: 306–321.
23. Ritterband LM, Thorndike FP, Cox DJ, et al. A behavior change model for internet interventions. *Ann Behav Med* 2009; 38: 18–27.. 2009/10/04.
24. Noar SM, Benac CN and Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007; 133: 673–693.
25. Krebs P, Prochaska JO and Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. *Prev Med* 2010; 51: 214–221.
26. Lustria MLA, Noar SM, Cortese J, et al. A meta-analysis of web-delivered tailored health behavior change interventions. *J Health Commun* 2013; 18: 1039–1069.
27. Valle CG, Queen TL, Martin BA, et al. Optimizing tailored communications for health risk assessment: a randomized factorial experiment of the effects of expectancy priming, autonomy support, and exemplification. *J Med Internet Res* 2018; 20: 63.
28. Muench F, van Stolk-Cooke K, Morgenstern J, et al. Understanding messaging preferences to inform development of mobile goal-directed behavioral interventions. *J Med Internet Res* 2014; 16: 14.. 2014/02/07.
29. Ambeba EJ, Lei Y, Sereika SM, et al. The use of mHealth to deliver tailored messages reduces reported energy and fat intake. *J Cardiovasc Nurs* 2015; 30: 35–43.
30. Bandura A and Locke EA. Negative self-efficacy and goal effects revisited. *J Appl Psychol* 2003; 88: 87–99.
31. Beck AT and Bredemeier K. A unified model of depression: integrating clinical, cognitive, biological, and evolutionary perspectives. *Clin Psychol Sci* 2016; 4: 596–619.
32. Eberly MB, Liu D, TR M, et al. Attributions and emotions as mediators and/or moderators in the goal-striving process. In: EA L and Latham GP (eds) *New developments in goal setting and task performance*. New York: Routledge, 2013, pp.35–50.
33. Weiner B. Attribution in personality psychology. In: Perrin LA (eds) *Handbook of personality: theory and research*. New York: Guilford Press, 1990, pp.465–485.
34. Heider F. *The psychology of interpersonal relations*. Hoboken, NJ, US: John Wiley & Sons Inc, 1958.
35. Peterson C, Semmel A, Von Baeyer C, et al. The attributional style questionnaire. *Cognit Ther Res* 1982; 6: 287–299.
36. Campbell CR and Martinko MJ. An integrative attributional perspective of empowerment and learned helplessness: a multimethod field study. *J Manage* 1998; 24: 173–200.
37. Abramson L, Metalsky G and Alloy L. The hopelessness theory of depression: a metatheoretical analysis with implications for psychopathology research. Manuscript submitted for publication, 1986.
38. Kleim B, Gonzalo D and Ehlers A. The depressive attributions questionnaire (DAQ): development of a short self-report measure of depressogenic attributions. *J Psychopathol Behav Assess* 2011; 33: 375–385.
39. Rubenstein LM, Freed RD, Shapero BG, et al. Cognitive attributions in depression: bridging the gap between research and clinical practice. *J Psychother Integr* 2016; 26: 103–115.
40. Tolli AP and Schmidt AM. The role of feedback, causal attributions, and self-efficacy in goal revision. *J Appl Psychol* 2008; 93: 692–701.
41. Sweeny K. Information avoidance: who, what, when, and why. *Rev Gen Psychol* 2010; 14: 340–353.
42. Hurley L, Nezami BT, Valle CG, et al. Motivated information avoidance in an mHealth weight loss intervention: associations between unmet behavioral goals and likelihood of viewing program messages. *DIGITAL HEALTH* 2024; 10: 20552076241287365.
43. Valle CG, Nezami BT and Tate DF. Designing in-app messages to nudge behavior change: lessons learned from a weight management app for young adults. *Organ Behav Hum Decis Process* 2020; 161: 95–101.
44. Liao P, Klasnja P, Tewari A, et al. Sample size calculations for micro-randomized trials in mHealth. *Stat Med* 2016; 35: 1944–1971.. 2015/12/28.
45. Klasnja P, Smith S, Seewald NJ, et al. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. *Ann Behav Med* 2019; 53: 573–582.
46. Smith SN, Lee AJ, Hall K, et al. Design lessons from a micro-randomized pilot study in Mobile health. In: Rehg JM, SA M and Kumar S (eds) *Mobile health: sensors, analytic methods, and applications*. Cham: Springer International Publishing, 2017, pp.59–82.

47. Bidargaddi N, Almirall D, Murphy S, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a Mobile health app. *JMIR Mhealth Uhealth* 2018; 6: e10123.. 2018/12/01.
48. Yardley L, Spring BJ, Riper H, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016; 51: 833–842.
49. Short CE, DeSmet A, Woods C, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. *J Med Internet Res* 2018; 20: e292.. 2018/11/18.
50. Perreira KM, Deeb-Sossa N, Harris KM, et al. What are we measuring? An evaluation of the CES-D across race/ethnicity and immigrant generation. *Social Forces* 2005; 83: 1567–1601.
51. Addis ME and Mahalik JR. Men, masculinity, and the contexts of help seeking. *Am Psychol* 2003; 58: 5–14.
52. Murphy SA and Almirall D. JOOL MRT Analysis, 2018.
53. Boruvka A, Almirall D, Witkiewitz K, et al. Assessing time-varying causal effect moderation in Mobile health. *J Am Stat Assoc* 2018; 113: 1112–1121.. 2018/11/24.
54. Qian T, Yoo H, Klasnja P, et al. Estimating time-varying causal excursion effects in mobile health with binary outcomes. *Biometrika* 2021; 108: 507–527.
55. Robins JM. Causal inference from complex longitudinal data. Latent variable modeling and applications to causality. Springer, 1997, pp.69–117.
56. Qian T. MRTAnalysisBinary, (2020, 2022).
57. Therneau TM and Grambsch PM. A Package for Survival Analysis in R. https://cran.r-project.org/package=survival (2022).
58. Weissman MM, Sholomskas D, Pottenger M, et al. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol* 1977; 106: 203–214.
59. Klasnja P, Hekler EB, Shiffman S, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. *Health Psychol* 2015; 34: 1220–1228.
60. Lyzwinski LN, Caffery LJ, Bambling M, et al. Consumer perspectives on mHealth for weight loss: a review of qualitative studies. *J Telemed Telecare* 2018; 24: 290–302.
61. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013; 46: 81–95.
62. Cooper S. An examination of some temporal implications of goal setting. *Paper presented at SIOP* 2007.
63. Jourden FJ. *The influence of feedback framing on the self-regulatory mechanisms governing complex decision-making*. Stanford, CA: Stanford University, 1991.
64. Webb TL, Chang BPI and Benn Y. The ostrich problem': motivated avoidance or rejection of information about goal progress. *Soc Personal Psychol Compass* 2013; 7: 794–807.